FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS
COMENIUS UNIVERSITY BRATISLAVA

Calibration of term structure models

Dissertation Thesis

Alexandra Urbánová Csajková

Supervisor: Daniel Ševčovič

Bratislava 2007

Contents

Chapter 1

Introduction

In recent decades, new interest rate dependent securities have appeared, like bond futures, options on bonds, swaps etc., whose payoffs are dependent on the interest rates. It means that their values are influenced by the interest rate. Expansion of the market with interest rate dependent securities, due to the need of risk management and also the speculative business, leads to valuation of these securities. In the construction of valuation models, it is important to incorporate the stochastic movement of interest rate into consideration. In this construction the random fluctuations of the interest rate market could be the right and realistic way. The stochasticity of the interest rate, especially the term structure of interest rate has to be modeled correctly.

The term structure is a functional dependence between the time to maturity of a bond and its yield. Relevant interest rate models characterize the bond prices (or yields) as a function of time to maturity, state variables e.g. instantaneous interest rate as well as several model parameters. Although several approaches for pricing interest rate derivatives have been proposed, no definite pricing model has been reached with regard to the best approach for these problems. Term structure models, as an important part in financial derivatives theory, have attracted a lot of attention from both a theoretical as well as practical point of view.

Much effort is being spent to calibrate interest rate models. Well known methods from the literature are for example the Generalized Method of Moments, the Gaussian estimation methods, the Monte Carlo filtering approach or the MCMC (Markov Chain Monte Carlo) method. These methods have been applied not only for bond pricing but also for interest rate swaps. Recently, other estimation methods for interest rate models have been proposed. These methods are based on other interest rate derivatives like e.g. prices of caps and floors. However, such derivatives are still not available in some financial markets including most of transitional Central European countries.

However, less attention is put on possible applications of interest rate models to Central European financial markets like Czech Republic, Slovakia, Poland and Hungary. Furthermore, a comparison to stable Western European financial markets has not been done yet.

There is a difference between Central and Western European financial markets. The convergence of former socialist countries and their financial markets to the western ones is evident. The question is: how far they are form each other? In another way: is the interest rate modeling different in these countries or not? To get an answer to this question we have to know the background of Central European markets and also to be familiar with the term structure modeling. Chapter 2 introduces the motivation for the calibration of interest rate models, the Cox-Ingersoll-Ross (CIR) and Vašiček interest rate models. We recall the term structure of interest rate and briefly review the basic properties of the one factor interest rate models, in particular CIR and Vašiček models.

There exist three main types of term structure models. In Chapter 3 there is a brief review of the equilibrium models, the no-arbitrage models, the LIBOR market models and the multi factor versions of some of them. First, the equilibrium models: they start with some assumptions about economic variables and imply a process for short-term interest rate. Second, the no-arbitrage interest rate models: they are designed to be consistent with today's term structure. In this class of models, the term structure is an input to the parameter estimation. Third, the LIBOR market models: these models present a tool for exotic interest rate derivatives pricing. They provide conditions on the drift of the forward rates if arbitrage is to be prevented. The above mentioned models are used in the pricing of interest rate derivatives. For some of them there exists an analytical solution. Unfortunately, for most of them, no analytical solution exists, neither exact nor approximate. The parameters of these models need to be specified properly for appropriate pricing of securities.

The most important estimation techniques and calibration methodologies, such as the Markov Chain Monte Carlo method, Generalized Method of Moments and the Maximum Likelihood Method are described, to compare their advantages and disadvantages to the developed and in this thesis presented new method. It tries to eliminate the disadvantages of the other methods.

Models which are analyzed in this thesis are the well known Cox-Ingersoll-Ross one factor interest rate model and Vašiček interest rate model. They belong to the set of equilibrium models. These models generate predicted term structures whose shape depends on the models parameters and the initial short rate.

Chapter 4 is focused on our goals. It contains topics on which is our work focused on.

The transformation of parameters of the CIR and Vašiček models and the optimal choice of some of these parameters are presented in Chapter 5. Using the introduction of new parameters it is possible to reduce the parameter space. This knowledge will be used during the calibration of the loss functional to real data. The loss functional need not be necessarily convex, so we have to introduce a robust numerical method for the optimization. This method is based on a variant of the evolution strategy. Of course there exist other methods solving the non-linear optimization problem. They are also discussed in this chapter. The calibration of two factor models is complex because of the involved second stochastic factor. We propose a suggestion on the loss functional for these models and also sketch its possible minimization.

The following part of the thesis consists of two main chapters where the internal and external methods of calibration are described. The main difference between these two methods is that we use extra information e.g. from an expert in the external calibration method. This is not used in the internal calibration method.

The internal calibration described in Chapter 6 is based on the two step optimization method. In the first step we find a global minimum of the non-linear loss functional and after that in the second step we find a maximum of the likelihood function over one dimensional curve consisting of global minimizers of the loss functional. This method uses information only from the data basis. The quality of the fit is measured by non-linear R^2 ratio. This calibration method uses only the means and the covariances from the data. It means that the whole term structure is not used in the calibration. Bounding of means or the term structures of interest rates uses more information from the data. This step is a connection between the internal and external calibration method. The application of the above mentioned internal calibration method is presented at the end of this chapter. This estimation method is extensively tested for several European financial markets, on real market term structures, the Euro-zone term structures as well as Central European financial markets. The binding interval approach is also tested on real data.

Comparing of the option pricing with the zero-coupon bond pricing and interest rate (yield curve), we can conclude that there exists an analogy between them. It means that not all parameters in the option pricing theory could be calibrated internally. This motivates us to try a new approach, the so-called external calibration methodology, utilizing an externally provided parameter. The prescribed expected long-term interest rate interval seems to be an appropriate choice. It is described in Chapter 7.

Discussion and concluding remarks are presented in Chapter 8.

The thesis will be focused on parameter calibration of term structure models. It will be preceded by parameter reduction and transformation. We will propose the two step method of reduced parameter identification (the loss functional minimization) in the first step and the transformed parameter identification in the second step. The calibration will be based either on internal methodology (the restricted maximum likelihood method or the mean value binding approach using the data basis) or on external methodology (using externally provided parameter or targeting interval). Some of these results have been published in papers [64, 65].

Chapter 2

Term structures and their modeling

In the most former socialist countries financial markets represented by banks and other financial institutions as well as capital markets did not exist [75]. Stock markets that existed before, e.g. the Warsaw Stock Exchange or the Prague Stock Exchange were completely closed during socialism. In these countries financial systems were based on one single institution, the so-called monobank. This bank was responsible for monetary policy and commercial banking. From the beginning of the integration process of these countries into the world economy and the European Union (EU) their financial sectors have developed and changed in two last decades.

The first few years of transition were mainly characterized by instability and restructuring. The monobank system has been changed into two-tier banking system meaning the separation of central bank and commercial bank functions. During this period the capital markets (especially the bond markets and stock markets) and money markets have been set up.

The second transition period was characterized by overall strengthening, development, and macroeconomic stability with positive economic growth rates. Financial sectors in the new EU member states are characterized by overall financial stability and a trend of positive financial development such as increase in the size and efficiency of the financial sector. The key features of these financial markets are the low level of intermediation, the strong dominance of the banking sector, strong presence of foreign owners and investors and the rise of institutional investments. The bond markets are generally small compared to western markets and dominated by government securities. The most developed bond markets are in Hungary, Poland and the Czech Republic. Derivatives markets appear to be quite active also in these countries.

Development of these financial markets is extremely dynamic. Interest rate derivatives became very popular. The main reason is that the investors need to be hedged against the interest rate risk which is the implication of unexpected interest rate changing or twisting the yield curve. So, there is a requirement for a robust interest rate derivative pricing model and a calibration methodology for this model to real market data. There has been done a lot of research related to the calibration of various types of interest rate models. They are mostly applications on stable Western European financial markets data. Less attention is put on the investigation of the Central European markets (the Czech Republic, Slovakia, Poland and Hungary) and their comparison to the stable western markets. There is a partial progress in this direction made by Vojtek in [74].

2.1 What are term structures?

A bond (cf. [47]) is a contract under which the borrower promises to pay the bondholder periodic coupon interest payments and par value on specific dates. If there is no coupon payment, the bond is said to be a zero-coupon bond. The par value is also called the face value of the bond or principal. This value must be repaid at the maturity date of a bond. The value of a bond is the present value of the cash flows which are realized during the life of the bond. Bonds are traded securities and their prices are observed in the market. The price of a bond depends on different factors, like the fluctuations in interest rates or the outstanding coupon payments.

Let $P(t,T)$ denote the price at time t of a zero-coupon bond maturing at time T. Par value is assumed to be $P(T,T) = 1$. Yield to maturity $R(t,T)$ is defined by:

$$R(t,T) = -\frac{1}{T-t}\ln P(t,T), \tag{2.1}$$

which gives the internal rate of return at time t on the bond. The yield curve is the dependence of $R(t,T)$ on T. Term structure of interest rate is a functional dependence between the yield and the time to maturity $T-t$ of a bond.

In the forward contract the holder agrees to buy one zero-coupon bond at a later time T_1 with maturity $T_2 > T_1$. The forward rate $f(t,T_1,T_2)$ is the rate in time t for the period between T_1 and T_2:

$$f(t,T_1,T_2) = -\frac{1}{T_2-T_1}\ln\frac{P(t,T_2)}{P(t,T_1)}. \tag{2.2}$$

The instantaneous forward rate is:

$$F(t,T) = -\frac{1}{P(t,T)}\frac{\partial P(t,T)}{\partial T}. \tag{2.3}$$

Integration of the above expression together with (2.1) gives the following equation:

$$R(t,T) = \frac{1}{T-t}\int_t^T F(t,u)du. \tag{2.4}$$

This indicates that the yield (or the bond price respectively) can be recovered from the term structure of the forward rates.

Overnight or the spot interest rate is the initial point of the yield curve and is defined as follows:

$$R(t,t) = F(t,t) = r_t = -\frac{\partial \ln P(t,T)}{\partial T}\bigg|_{t=T}. \tag{2.5}$$

Several theories of term structures have been proposed to explain the shape of a yield curve. First, the expectation theory stating that the long-term interest rates reflect expected future short-term interest rate. Second one is the market segmentation theory stating that the yield curve will depend on the supply and demand conditions for funds. Third one is the liquidity preference theory stating that the lenders prefer to make short-term loans ([47]).

The analysis of term structure is important in the analysis of interest rate derivatives and their pricing models, so it becomes also crucial in the calibration methodology.

2.2 Modeling interest rate fluctuations by means of Wiener processes

A Wiener process is a special type of Markovian stochastic process. It is a stochastic process describing the probabilistic evolution of the value of a variable through time and because of the Markovian property, only the present value of the variable is relevant for the future predicting. The past values and also the way how the variable emerged from the past is irrelevant. These features are summarized in the following definition.

Definition 2.2.1. *The standard Wiener process $w(t), t \geq 0$ is a stochastic process with the following properties:*

- *every increment $w(t+\Delta) - w(t)$ is normally distributed with mean 0 and variance $\Delta > 0$, for any $t \geq 0$,*
- *for every $0 < t_1 < ... < t_n$ increments $w(t_2) - w(t_1), ..., w(t_n) - w(t_{n-1})$ are independent random variables with distribution given in the first point,*
- *$w(0) = 0$ and the sample paths of $w(t)$ are continuous.*

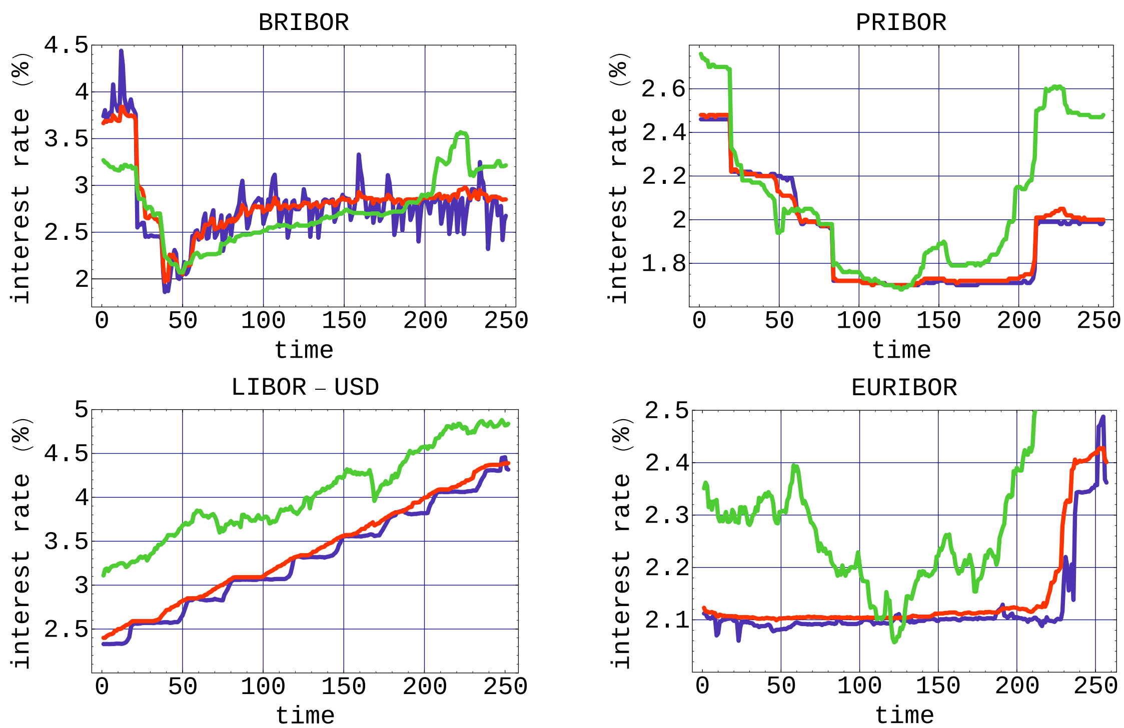


Figure 2.1: Trend of interest rates in 2005 for BRIBOR, PRIBOR, USD-LIBOR and EURIBOR with different maturities (1 week - blue, 1 month - red, 1 year - green). Horizontal axis represents time in days.

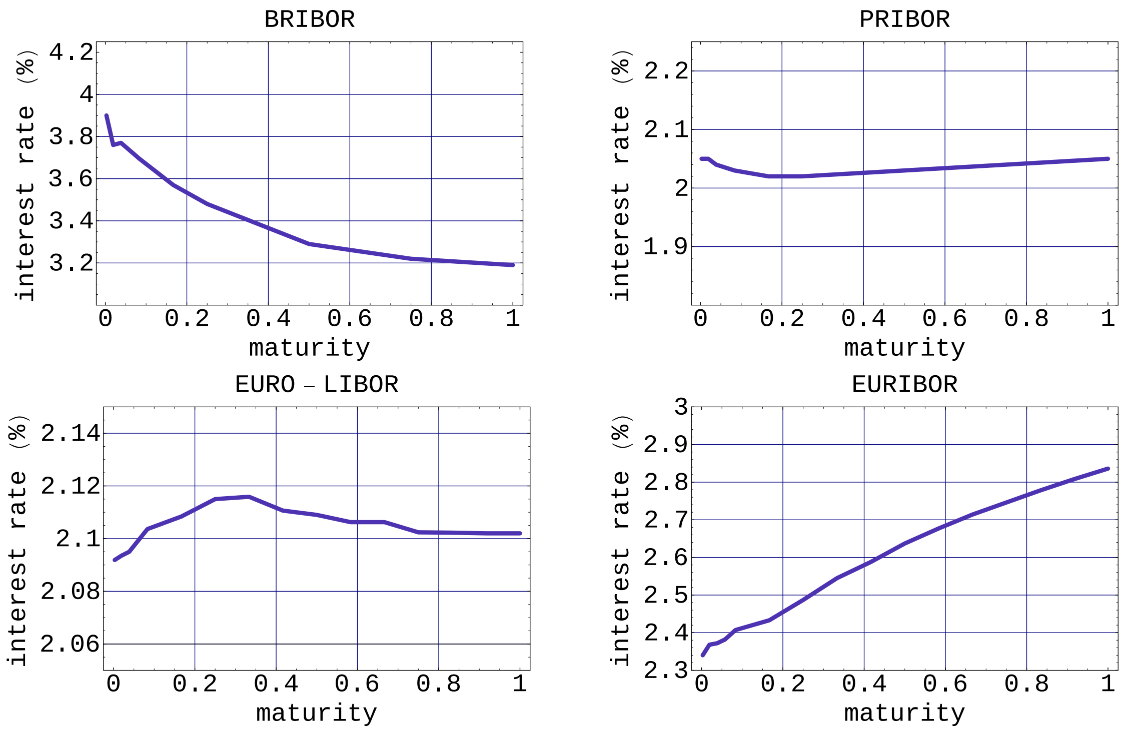


Figure 2.2: Term structures of interest rates for BRIBOR, PRIBOR, EURO-LIBOR and EURIBOR. Horizontal axis represents maturity in years.

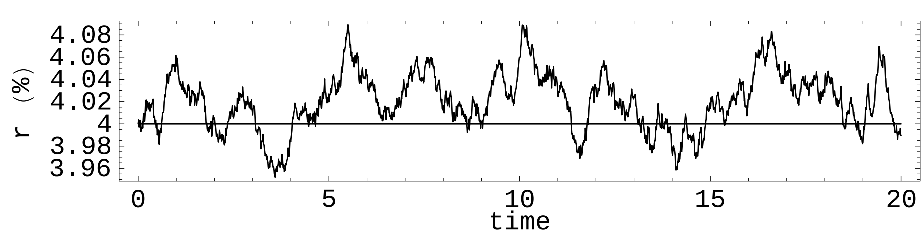


Figure 2.3: Simulation of a mean reverting process driven by (2.6) with parameters: $\kappa = 1, \sigma = 0.2, \theta = 0.04, \gamma = 0.5$ within the time interval $(0, 20)$.

2.3 One factor interest rate models

The basis of the term structure modeling is the stochastic process of the short-term interest rate in risk-less world and its influence on the bonds. It is important to realize that there does not exist a process truly describing the interest rate fluctuations. As we can see on Figure 2.1, the shapes of interest rates are different for various financial markets. They also change by the time to maturity, but the trend is the same for the short and the long one. Hence the modeling of such a different and complicated shape of interest rates is difficult. In many models the basic assumption is that the bond price depends on only one stochastic variable. These models are called the one factor bond pricing models.

Let us assume that the short/spot rate r_t [1] follows a special type of continuous Markovian stochastic process, the mean reverting process of the form:

$$dr_t = \kappa(\theta - r_t)dt + \sigma r_t^{\gamma} dw_t, \tag{2.6}$$

where $\{w_t, t \geq 0\}$ denotes the standard Wiener process, κ, θ, σ are positive constants and γ is a nonnegative constant.

Parameter κ is the speed of reversion, σ is the volatility of the process and θ is the limiting interest rate. The parameter γ is determining the type of the model. If $\gamma = \frac{1}{2}$ than the model derived from (2.6) is refereed to as the Cox-Ingersoll-Ross (CIR) model. If $\gamma = 0$ then it is called the Vašiček model. Besides these two models there exist a spectrum of one factor interest rate models, e.g. Dothan model, Brennan and Schwartz model, etc.

In Figure 2.3 we plot sample data obtained from a simulation of equation (2.6) to demonstrate how the Wiener process works in the interest rate modeling.

Using equation (2.1), the whole term structure can be determined as a function of r_t once κ, θ, σ and γ have been chosen in (2.6). The shape can be upward sloping, downward sloping, or slightly humped as we can see on Figure 2.2 which is based on real market data.

The main point of deriving a one factor model is the following Proposition which can be considered as the extension of the rules of differential in ordinary calculus to stochastic calculus.

Proposition 2.3.1. *(Itô lemma) ([47], p. 29) Let $f(x,t)$ be a C^2 smooth, non-random function and $x(t)$ a stochastic process defined by:*

$$dx = \mu(x,t)dt + \sigma(x,t)dw$$

where w is the standard Wiener process. Then the stochastic process $f(x(t),t)$ satisfies the following stochastic differential:

$$df = \left(\frac{\partial f}{\partial t} + \mu(x,t)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 f}{\partial x^2} \right) dt + \sigma(x,t)\frac{\partial f}{\partial x} dw.$$

[1] From now on we will denote the short/spot rate $r_t = r$.

We assume that the short rate follows the mean reverting process (2.6). According to Itô lemma the differential for the bond price $P = P(t, T, r)$ [2] is the following:

$$\begin{aligned} dP &= \left(\frac{\partial P}{\partial t} + \mu_r \frac{\partial P}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 P}{\partial r^2} \right) dt + \sigma_r \frac{\partial P}{\partial r} dw \\ &=: \widetilde{\mu}(t,T)dt + \widetilde{\sigma}(t,T)dw \end{aligned} \tag{2.7}$$

where $\mu_r = \kappa(\theta - r_t)$ and $\sigma_r = \sigma r_t^\gamma$. The $\widetilde{\mu}(t,T)$ is the drift rate and $\widetilde{\sigma}(t,T)$ is the variance rate of the stochastic process of P. To hedge the bond we construct a portfolio of two bonds with different maturities. The interest rate can not be used in the hedging because it is not a traded security. The portfolio is constructed from one bond with maturity T_1 and Δ bonds with maturity T_2:

$$\pi = P_1(t, T_1, r) + \Delta P_2(t, T_2, r). \tag{2.8}$$

The investor determines the Δ to minimize his risk. Change in portfolio value in time dt is:

$$d\pi = dP_1 + \Delta dP_2.$$

According to the bond price dynamics defined by (2.7) we get:

$$d\pi = (\widetilde{\mu}_1 + \Delta\widetilde{\mu}_2)dt + (\widetilde{\sigma}_1 + \Delta\widetilde{\sigma}_2)dw. \tag{2.9}$$

To eliminate the risk in the portfolio we put $\widetilde{\sigma}_1 + \Delta\widetilde{\sigma}_2 = 0$. So we get the proportion of the bonds with different maturities $\Delta = -\frac{\widetilde{\sigma}_1}{\widetilde{\sigma}_2}$, and the value of the hedged portfolio is

$$d\pi = \left(\widetilde{\mu}_1 - \frac{\widetilde{\sigma}_1}{\widetilde{\sigma}_2}\widetilde{\mu}_2 \right) dt. \tag{2.10}$$

Since the portfolio is instantaneously risk-less, it must earn the risk-less spot interest rate that is $d\pi = r\pi dt$. Combining these two results for the portfolio change in time dt we obtain:

$$\frac{rP_1 - \widetilde{\mu}_1}{\widetilde{\sigma}_1} = \frac{rP_2 - \widetilde{\mu}_2}{\widetilde{\sigma}_2}.$$

This ratio is valid for all maturities, so the relation

$$\widetilde{\lambda}(r,t) = \frac{\widetilde{\mu}(t,T) - r(t)P(t,T)}{\widetilde{\sigma}(t,T)} \tag{2.11}$$

is independent of maturity T. The quantity $\widetilde{\lambda}(r,t)$ is called the market price of risk, since it gives the extra increase in expected instantaneous rate of return of bond per an additional unit of risk. Dependence of the bond on the investors preferences could

[2] From now on we will denote the price of zero-coupon bond $P = P(t,T,r) = P(\tau, r)$.

not be eliminated as it is possible in the stock/option hedge, because the interest rate is not a traded security. If we combine the equation (2.7) with (2.11) we obtain:

$$\begin{aligned} \widetilde{\sigma} &= \sigma_r \frac{\partial P}{\partial r} \\ \widetilde{\mu} &= \widetilde{\lambda}\sigma_r \frac{\partial P}{\partial r} + rP = \frac{\partial P}{\partial t} + \mu_r \frac{\partial P}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 P}{\partial r^2}. \end{aligned}$$

Proposition 2.3.2. *([47], p. 321) The governing equation for the price of the zero-coupon bond $P = P(t, T, r)$ is a parabolic partial differential equation of the form:*

$$\frac{\partial P}{\partial t} + (\kappa(\theta - r) - \widetilde{\lambda}\sigma r^{\gamma})\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 P}{\partial r^2} - rP = 0, \tag{2.12}$$

where $t \in (0, T)$ and $r > 0$, satisfying $P(T, T, r) = 1\ \forall r$.

The parameter $\widetilde{\lambda}$ is different for the two discussed models. For the CIR model we take $\widetilde{\lambda}(r) = \lambda r^{\frac{1}{2}}/\sigma$ whereas for Vašiček model we take $\widetilde{\lambda}(r) = \lambda$, where λ is a constant.

Proposition 2.3.3. *([47], pp. 322-325) There exists an explicit solution of PDE (Partial Differential Equation) (2.12) for both models and it is of the form:*

$$P(T - \tau, T, r) = A(\tau)e^{-B(\tau)r}, \tag{2.13}$$

where $\tau = T - t \in [0, T]$ and the functions $A(\tau), B(\tau)$ satisfy

$$\begin{gathered} B(\tau) = \frac{1 - e^{-\kappa\tau}}{\kappa}, \\ A(\tau) = \exp\left[(B(\tau) - \tau)\left(\theta - \frac{\sigma^2}{2\kappa^2} - \frac{\sigma\lambda}{\kappa}\right) - \frac{\sigma^2 B(\tau)^2}{4\kappa}\right], \end{gathered} \tag{2.14}$$

for the Vašiček model, and

$$\begin{gathered} B(\tau) = \frac{2(e^{\eta\tau} - 1)}{(\kappa + \lambda + \eta)(e^{\eta\tau} - 1) + 2\eta}, \\ A(\tau) = \left(\frac{\eta e^{(\kappa+\lambda+\eta)\tau/2}}{e^{\eta\tau} - 1} B(\tau)\right)^{\frac{2\kappa\theta}{\sigma^2}}, \end{gathered} \tag{2.15}$$

for the CIR model, where $\eta = \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}$.

Note that the market price of risk appears only in summation with κ, as it was first pointed out in paper [56] by Pearson and Sun.

The main difference between the Vašiček and CIR model is that in the case of Vašiček model the interest rates may become negative. This negative feature of this model is eliminated in the CIR model. Since $P(t, T, r)$ depends only on the difference $\tau = T - t$ we shall henceforth write $P = P(\tau, r) = P(T - \tau, r)$.

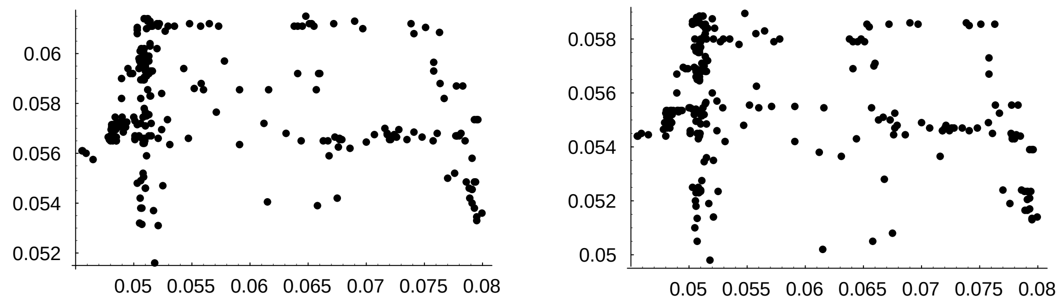


Figure 2.4: The overnight interest rate against the interest rate for longer time horizon; 6 month (left) one year (right)

2.4 Multi factor models

In the above one factor models, the whole term structure was dependent on one stochastic process. The advantages of these models are the analytic tractability and simplicity. In many times closed form solutions for bonds and term structures exist. On the other hand this approach tends to oversimplify the interest rate process. Multi factor models provide better solution to this problem. However, the analytic tractability becomes complicated.

The most popular multi factor models are two factor models. The first class of these models uses the short and long rate as state variables, while the second class uses the short rate and its variance as state variables. Let us mention only some examples of the second class: the Schaefer and Schwartz model [47], the Fong and Vašiček model (see [8, 33, 47, 66, 69]), the Chen and Scott model [22], the Longstaff and Schwartz model [47], the Anderson and Lund [47] and the two factor CIR model.

The main step of deriving a two factor models is a construction of risk-less portfolio of three bonds with different maturities and applying the multi-dimensional Itô lemma [47]. The implementation of these models is cumbersome and that is criticized by practitioners.

The price of a bond in the one factor model is a function of the spot rate if the time to maturity is given. It means that to each value of the spot rate one yield curve is assigned. The multi factor models allow to assign different yield curves to the same short rate depending on the values of the other factors.

Demonstration of this feature has been shown by Stehlíková [68] on BRIBOR's yield curves. The plot of the overnight interest rate (the approximation of the spot rate) against the interest rate for longer time horizon T (but on the same day t) shows that for the same values of the overnight the value of the other assigned rate are from a bigger interval. This indicates that the usage of the multi factor models is necessary and reasonable.

Let us assume that in the two factor interest rate model, the short rate r and its volatility y are the two stochastic factors and they follow a system of stochastic differential equations:

$$
\begin{aligned}
dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}r^\gamma dw_1, & (2.16)\\
dy &= \kappa_2(\theta_2 - y)dt + vy^\delta dw_2. & (2.17)
\end{aligned}
$$

Constants $\theta_1 > 0$ and $\theta_2 > 0$ are the limiting interest rate and limiting dispersion, respectively; $\kappa_1 > 0$ and $\kappa_2 > 0$ are speed of reversions for the short rate and volatility of it and w_1, w_2 are Wiener processes (c.f. [47]) with correlation $\rho \in [-1, 1]$, $\gamma, \delta \geq 0$ are model parameters and $v > 0$ is the volatility of the short rate volatility.

As it is depicted in [69], $y(t) \to \sigma_1^2 := \theta_2$ as $t \to \infty$ in the case $v = 0$. It means that the form of governing equation for the short rate reduces to $dr = \kappa_1(\theta_1 - r)dt + \sigma_1 r^\gamma dw_1$. As we can see it is the same equation as for the short rate in the one factor model case. Now we can turn to the two factor model.

Proposition 2.4.1. *([47], pp.329-330) The governing equation in the two factor case for the price of the zero-coupon bond $P = P(\tau, r, y)$, where $\tau = T - t$, is a partial differential equation of the form:*

$$
\begin{aligned}
&-\frac{\partial P}{\partial \tau} + (\kappa_1(\theta_1 - r) - \tilde{\lambda}_1\sqrt{y}r^\gamma)\frac{\partial P}{\partial r} + (\kappa_2(\theta_2 - y) - \tilde{\lambda}_2 v y^\delta)\frac{\partial P}{\partial y} + \\
&\frac{1}{2}(\sqrt{y}r^\gamma)^2\frac{\partial^2 P}{\partial r^2} + \frac{1}{2}(vy^\delta)^2\frac{\partial^2 P}{\partial y^2} + (\sqrt{y}r^\gamma)(vy^\delta)\rho\frac{\partial^2 P}{\partial r \partial y} - rP = 0 \qquad (2.18)
\end{aligned}
$$

where $\{(r, y), 0 \leq r, 0 \leq y\}$ with the initial condition $P(0, r, y) = 1$.

In this model $\tilde{\lambda}_1$ represents the so-called market price of risk whereas $\tilde{\lambda}_2$ is the market price of volatility (cf. [33]).

If $\gamma = 0$ and $\delta = \frac{1}{2}$ then we get the so-called Fong-Vašiček model derived by Fong and Vašiček in 1991 (see also [8, 33, 66]). In this model, the market prices of risk and volatility are as follows:

$$
\tilde{\lambda}_1 = \lambda_1\sqrt{y} \quad \& \quad \tilde{\lambda}_2 = \lambda_2\sqrt{y}
$$

where λ_1 and λ_2 are constants.

A solution to the PDE (2.18) for the Fong-Vašiček model is of the form:

$$
P(\tau, r, y) = A(\tau)e^{-B(\tau)r - C(\tau)y} \qquad (2.19)
$$

iff $A = A(\tau),\ B = B(\tau), C = C(\tau), \tau \in (0, T]$, satisfy the following system of ordinary differential equations:

$$
\begin{aligned}
\dot{A} &= -A\left(\kappa_1\theta_1 B + \kappa_2\theta_2 C\right), \\
\dot{B} &= -\kappa_1 B + 1\,, \qquad (2.20)\\
\dot{C} &= -\lambda_1 B - \kappa_2 C - \lambda_2 v C - \frac{B^2}{2} - \frac{v^2C^2}{2} - v\rho BC
\end{aligned}
$$

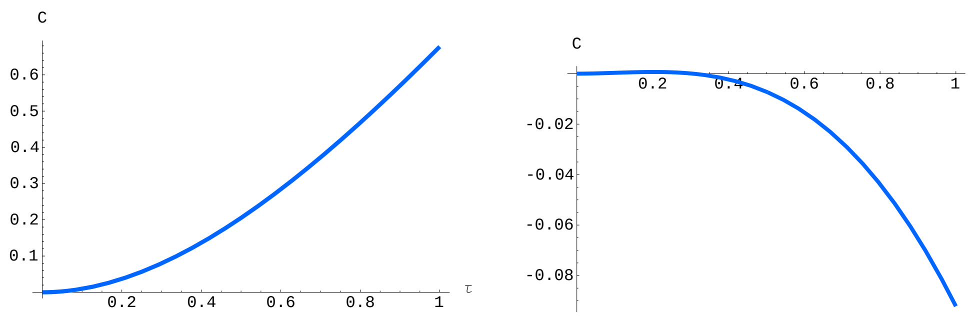


Figure 2.5: Plots of the function C for two different sets of parameters; $\lambda_1 = -2, \kappa_1 = 0.5$ (left) and $\lambda_1 = -0.1, \kappa_1 = 0.2$ (right). In both cases we chose $\lambda_2 = -3, \kappa_2 = 0.2, \theta_1 = 0.04, \theta_2 = 0.2, v = 0.1, \rho = 0.5$.

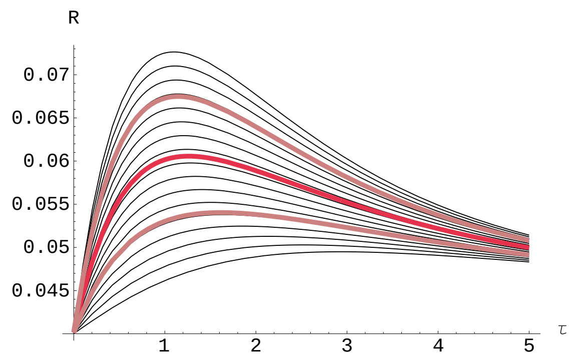


Figure 2.6: Yield curve $R = R(\tau, r, y), \tau \in [0, T]$, for several values of the y variable. The mean value of y is plotted in the middle in red. The 95% confidence interval of the term structure R is bounded by upper and lower pink curves.

with initial conditions $A(0) = 1,\ B(0) = 0,\ C(0) = 0$. We denoted by $(\dot{})$ the derivation $\partial_\tau(\ )$. Here we assume that $\theta_2 = \sigma_1^2$. The system of ODEs (Ordinary Differential Equations) (2.20) has been derived by Stehlíková and Ševčovič [69].

We can find an analogy with the one factor version of the Vašiček model. Functions $A(\tau)$ and $B(\tau)$ satisfy the same system of differential equations if we choose $\widetilde{\lambda}(r) = \lambda_1\sigma_1$ in (2.12). $A(\tau)$ and $B(\tau)$ can be expressed as in (2.14). The behavior of $C(\tau)$ is depicted on Figure 2.5 (for more details see [69]). Yield curves derived from the solution to the PDE (2.18) for the Fong-Vašiček model for varying $y = \sigma_1^2$ is demonstrated on Figure 2.6.

Chapter 3

Survey of known estimation and calibration methods

The interest rate instruments become very popular and traded on both the exchange as well as over-the-counter that suit the specific demands of individual investors. The most known are the bonds and the options, but there exist other interest rate derivative products, like swaps, swaptions, interest rate caps and floors.

The interest rate swap (cf. [42, 47]) is an agreement to exchange interest rate payments for a fixed time period. In the case of plain vanilla interest rate swap two parties exchange cash flows equal to interest at fixed rate r^* and floating rate r on the same notional principal for a predetermined period of time. Exactly they exchange only the net difference in the cash flow payments on regular basis. To avoid the credit risk from this transaction an intermediary is used. It is easy to derive the governing equation for this plain vanilla interest rate swap. We assume continuous exchange of cash flows. So let $W(r,t)$ be the value of the swap. The swap can be understand that one party recieves coupon payment at rate $r - r^*$ on a simple bond with zero par value at the swap maturity date T_S. Suppose that the short rate process is a diffusion process of the form:

$$dr = \mu(r,t)dt + \sigma(r,t)dw$$

(general form of the equation (2.6)) then the governing equation for the $W(r,t)$ will be similar to the zero-coupon bond equation (2.12) except the coupon payment term $r - r^*$, i.e.

$$\frac{\partial W}{\partial t} + (\mu - \lambda\sigma)\frac{\partial W}{\partial r} + \frac{1}{2}\sigma^2\frac{\partial^2 W}{\partial r^2} - rW + r - r^* = 0. \tag{3.1}$$

We also need a terminal condition $W(r, T_S) = 0$.

Swaption (cf. [42, 47]) is an option to enter the swap at some time in the future. The governing equation for the swaption is the same as for the zero-coupon bond. The feature of this contract is reflected in the final condition:

$$V(r,T) = \max(\eta(W(r,T) - X), 0),$$

where η is a binary variable which takes the value 1 for a call and -1 for a put, $W(r,T)$ is the value of the swaption at expiration and X is the strike price of the option in the swaption. A cap (cf. [42, 47]) guarantees that the interest rate charged on a floating rate loan at any given time will be the minimum of the prevailing and ceiling rates. A floor (cf. [42, 47]) is the opposite to a cap. Interest rate cap(floor) can be considered as a portfolio of European put/call options on discount bonds.
In the pricing of these interest rate dependent securities, interest rate has a key role. This is the main reason why interest rate models are so popular. They try to capture the stochasticity of the interest rate.

In a deterministic world, an interest rate is fully determined and the value of a bond can be easily calculated. This consideration is very convenient but is not realistic for longer time horizons. Interest rates and bond prices changes are not deterministic and smooth. Hence there is a need to model interest rates stochastically.

There are essentially three generations of interest rate models. The first generation is modeling the interest rate directly. They are the so called short-term interest rate models. Models of the second generation comprise the Heath-Jarrow-Morton (HJM) models modeling the entire forward curve. The HJM type approach automatically fits the yield curve. The driving state variable of this model is the forward rate. It can be shown that all short rate model can be formulated in the HJM framework [47]. The last generation comprises the LIBOR market model (LMM) or Brace-Gatarek-Musiela model which attempts to model specific parts of the forward curve.

The above mentioned first generation of interest rate modeling has been used in pricing of interest rate derivatives in a number of different ways. One of the oldest approaches is based on modeling the short-term interest rate by Merton [47] and by Vašiček [42]. The main assumption of these works is the normality of the interest rates, so there is a possibility to become negative. Dothan [47], Rendleman and Bartter [59] proposed a log-normal distribution for the short-term interest rate and have avoided this disadvantage. Cox, Ingersoll, and Ross [42] proposed instead a non-central χ^2 distribution. Another example of these models is the Brennan-Schwartz [47] model. The above mentioned models are endogenous term structure models. Hence the initial term structure is the output of the model. In addition, the general equilibrium conditions are used to endogenize the interest rate and the price of all contingent claims.

The problem of the first generation is that they in general do not fit the initial yield curve. This problem have been solved by the second generation of the interest rate models, the so called no-arbitrage models. In this set of models the initial term structures are taken as inputs to the model and the values of contingent claims obtained from them are automatically consistent with these inputs. The main representatives of this generation are Ho and Lee [42] model, the Hull and White [42] model which is an extension of the CIR model or the Vašiček model, the Black-Derman-Toy (BDT) model [13], the Black and Karasinski [47] model and also the HJM [47] model. Hull and White model can be characterized as the Ho and Lee model with special selection of its parameters. The BDT model is similar to Ho and Lee model. Some of these models have good analytic tractability (like the extended Vašiček model), while others may be non-Markovian in nature (like the HJM model). The non-Markovian models can become less used in practice.

The last generation of interest rate modeling is the LIBOR market model. The LMM is industry standard model for pricing interest rate derivatives and is based on HJM forward rate approach. Assuming a conditional log-normal process for LIBOR, it builds a process for LIBOR interest rate. Many implementations of this model use Monte Carlo simulation to price European-style and Bermudian-style swaptions. This generation of models is used to calibrate cap and swaption volatilities. Most recent works dealt in fact with forward LIBOR or swap rates, e.g. Miltersten, Sandmann and Sondermann [53], Brace, Gatarek and Musiela [15] and Jamshidian [44].

In the above mentioned so called one factor interest rate models, the short-term interest rate is assumed to follow a one factor continuous time process. Most of these models have closed form solution, it means that the term structures and bond prices can be calculated explicitly. As we know, this approach tends to oversimplify the behavior of interest rate movements, so there is a research concentrated on the construction of multi factor models. The analytic tractability become more complicated and in most cases numerical methods for the evaluation of bond prices have to be used.

Recently, there has been a development in the multi factor models framework e.g. the Chen [21] three factor model, the multi factor version of the Vašiček model presented by Babbs and Nowman [5] and the multi factor equilibrium model in the CIR framework developed by Chen and Scott [22].

So far we have described two nodes of the Figure 3.1, the interest rate derivatives end the interest rate models. We also know why they are closely related through pricing. The last node, the estimation and calibration methods, is somehow specific. This node is close-knit with the other two. The reason of this is the following: if you have a market with different interest rate derivatives, you can calibrate them through their market prices and obtain the unknown parameters of the interest rate models included in their prices; or if you are on a market without any sophisticated interest rate dependent securities, you can estimate the parameters of the interest rate models through the interest rate observed on the market.

There are many attempts spent on calibration and estimation of these interest rate models whether through the interest rate or through the derivatives of interest rates. Let us introduce some examples of calibration and estimation techniques for the above mentioned models falling within the three main generations of interest rate models.

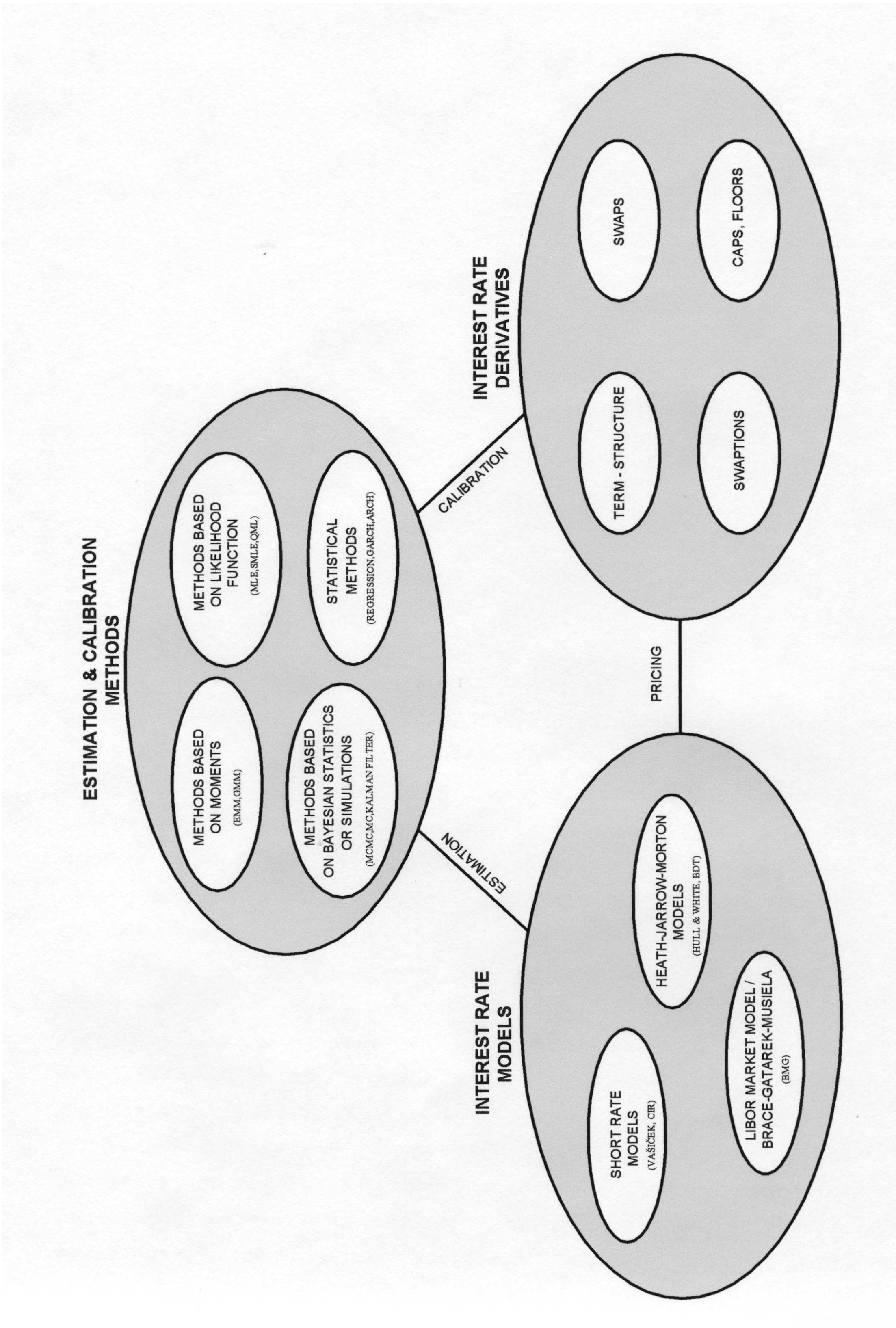


Figure 3.1: Schematic diagram of various calibration and estimation methods

- **1st generation:** Chan, Karolyi, Longstaff and Sandres [20] estimated and compared a variety of short-term interest rate models using the Generalized method of moments. They have determined which model best fits the short-term Treasury bill yield data.

  Later, Li [49] estimated one factor interest rate model parameters also by Generalized method of moments but in the Australian context. He showed that the unrestricted one factor model best fits the historical interest rate.

  Jensen [46] has demonstrated that the Longstaff and Schwartz model is inadequate for the description of the short-term interest rate, by a new implementation of the Efficient method of moment estimation principle. His main results are robust to sub-period analysis.

  Duan, Gauthier, Simonato and Zaanoun in [27] have applied another estimation method, the Maximum likelihood estimator, for the estimation of the Merton-Longstaff-Schwartz model parameters.

- **2nd generation:** Boyle, Tan and Tian [14] have investigated mathematical conditions under which fitting of another one factor interest rate model, the so called Black-Derman-Toy model, results in a reasonable calibration. The advantage of this model is that it can be calibrated to current market term structure of interest rates and of volatilities. Their main result is: if the current implied forward rates and the short rate volatilities are all positive, then the calibration of BDT model is possible.

- **3rd generation:** Takahashi and Sato [73] have developed a new methodology for estimation of general class of term structure models based on a Monte Carlo filtering approach. The method was applied to LIBORs and inter-bank rates swaps in the Japanese market.

  Vojtek [74] has presented a methodology to calibrate multi factor interest rate models for Central European countries. He has estimated the Brace, Gatarek and Musiela model parameters, especially the conditional volatilities and correlations, by a special type of GARCH model. The BGM model (also known as LIBOR market model) has been proposed and first time calibrated by Brace, Gatarek and Musiela [15], Jamshidian [45] and Miltersen, Sandmann and Sondermann [53].

- **multi factor models:** Pearson and Sun [56] have proposed an empirical method to estimate and test an extension of a two factor CIR model using data on discount and coupon bonds. Their result has shown failure of calibration based on Treasury bills. Chen and Scott [22] have presented a method for estimating multi factor version of the CIR model. The fixed parameters have been estimated by applying an approximate Maximum likelihood estimator using US Treasury market data and the unobservable factors have been estimated by non-linear Kalman filter.

As we can see several approaches have been used on the empirical estimation of interest rate models. They were based on the Generalized method of moments, Maximum likelihood estimation, Monte Carlo filtering or non-linear Kalman filter (a special type of the Markov Chain Monte Carlo method). Less attention is however put on one factor model parameters estimation and its possible application to Central European countries.

In the next sections we would like to present some of well known calibration and estimation methods and find out their advantages and disadvantages. The goal of this section is not to compare these methods, but learn from their pros and cons and propose a new method according to this knowledge.

3.1 Markov Chain Monte Carlo method

The Markov Chain Monte Carlo (MCMC) based methods are used for the continuous-time asset pricing models estimation. These methods are able to estimate equity price models with factors and multi factor term structure models with stochastic volatility. In the asset pricing models, the main goal is to get information about the state variables and the parameters from the asset prices. The solution to this problem is the distribution of the parameters, θ, and the state variables, X, conditional on observed prices, Y, denoted by $p(\theta, X|Y)$. Characterizing this distribution is difficult mainly due to the following reasons: (i) the model assume continuous-time, while the data are observed discretely; (ii) the $p(\theta, X|Y)$ is of very high dimension due to the state variable; (iii) mostly the distribution of the variables is non-normal and non-standard; (iv) in the case of term structure models the parameters enter in a non-linear way to PDE.

The first step in this approach is to interpret the continuous-time asset pricing model as state space model, in particular the non-linear, non-Gaussian type. The MCMC provides a general methodology in this case, and gives the distribution of the state variables and the parameters from the data. The Kalman filter approach is a special type of the MCMC method. It is the case when the state space model is linear and Gaussian with known parameters.

The MCMC is a simulation methodology, because it generates random samples from a given target distribution $p(\theta, X|Y)$. The theory of this method is based on the Hammersley-Clifford theorem [12]. It implies that the knowledge of $p(X|\theta, Y)$ and $p(\theta|X, Y)$ fully characterize the joint distribution $p(\theta, X|Y)$. It is much more easier to characterize the two conditional densities than the joint density. The algorithm generates a sequence of random variables called Markov Chain $\{X^{(g)}, \theta^{(g)}\}_{g=1}^{G}$ in the following way: consider two initial draws $X^{(0)}$ and $\theta^{(0)}$, then draw $X^{(1)} \sim p(X|\theta^{(0)}, Y)$ and $\theta^{(1)} \sim p(\theta|X^{(1)}, Y)$; and continue in the same manner.

If the two conditional densities are known in closed form and can be directly drawn from, the above mentioned algorithm is a Gibbs sampler. In other cases it is known as Metropolis-Hastings algorithm. The combination of these steps (Gibbs steps and the Metropolis-Hastings steps) generate the MCMC method.

Decomposition of the posterior $p(\theta, X|Y)$ into likelihood function $p(Y|X, \theta)$, distribution of the state variables $p(X|\theta)$ and prior distribution of the parameters $p(\theta)$ is based on the Bayes rule. If these distributions can be directly sampled using some standard method there is no problem with the algorithm. This is the case of the Gibbs sampler. In the case of the Metropolis-Hastings algorithm, the researcher has to specify a so called proposal density function because of no direct sampling of some of the conditional distributions. The choice of the proposal density will effect the convergence of the method. In extreme case the algorithm may never converge. Another problematic point of the MCMC method could be the prior distribution, especially using of non-informative priors.

The case of term structure models is a little bit complicated. The parameters enter the state space in non-linear often non-analytical fashion and also there are problems with stochastic singularities. Let us mention that in both the Vašiček and the CIR model the Metropolis algorithm has to be used, even more the CIR model involve conditional heteroscedasticity in the spot rates which can be solved with the heteroscedastic version of Kalman filter.

This approach has been used by number of researchers, some representatives of this line are: Carlin, Polson and Stoffer [19], Jacquier, Polson and Rossi [43], Fruhwirth-Schnatter and Geyer [34], Elerian, Chib and Shephard [31], Eraker [32].

3.1.1 Kalman filter approach

In the term structure analysis there are two main issues which have to be solved. One is modeling and estimation the current term structure of spot rates and the second is modeling and estimation of the dynamics of the term structure. The joint solving of this problem is achieved by the Kalman filter approach (KFA). The main advantage of the Kalman filter is that it uses all present and past price information to estimate the current term structure.

The KFA is based on a state-space representation of the term structure model suggested by Harvey [40] where the underlying state variable is treated as unobservable. In this formulation the observable interest rates are assumed to be related to unobservable state variables via a measurement equation and the unobservable state variables assumed to follow a Markov process described by a transition equation. So the aim of the Kalman filter is to obtain information about the unobservable state variables from the observed interest rates.

The vector of unobservable state variables is governed by stochastic differential equation and is defining the instantaneous interest rate. According to the form of the stochastic differential equation we can get several type of interest rate models. The value of the discount bond is obtained by applying the standard no-arbitrage assumptions. This implies that the KFA can be applied to a broad class of dynamic interest rate models including the CIR and Vašiček model.

Extension of this method can be used for models of interest rates using panel-data with missing observations which is quite common in many emerging markets. The model can be applied to value and hedge interest rate derivatives and estimate the term structure for days with small number of traded bonds.

Recent applications of this methodology to dynamic models of interest rates include: Lund [51, 52], Ball and Torous [7], Duan and Simonato [28], Geyer and Pichler [35], Duffe [29], de Jong and Santa-Clara [25], Babbs and Nowman [6], de Jong [24], Chen and Scott [22], Rossi [26] and Cortazar, Schwartz and Naranjo [23].

3.2 Maximum likelihood method

The Maximum likelihood method (MLM) is used in many papers in the current literature to estimate model parameters through the time-series approach, the cross-section approach or the combination of these two methods. The basis of this method is the maximum likelihood estimator which is the value that maximizes the (log) likelihood. This estimation method requires the knowledge of the state variable which is in the case of the one factor term structure models the instantaneous interest rate. If we can observe the state variable, we can estimate the model with the maximum likelihood method, because the likelihood function can be derived from the equation for the state variable. The closed form of the likelihood function is known only in some cases. For example the CIR model implies a non-central χ^2 and the geometric Brownian motion implies a log-normal density, for which closed forms of the likelihood functions are available. However, the closed form for some model is difficult to implement as it involves the Bessel function, and for some models no closed form is available. In these cases there are used only approximations of the densities, but this implies the disruption of the optimality properties of the estimator.

An important problem could occur in this method of calibration. This problem is about the existence and non-existence of the maximum of likelihood function (LF). An example of non-existence of the maximum of LF:

$$\ln L = -\frac{1}{2}\sum_{t=2}^{N}\left(\ln \nu_t^2 + \frac{\varepsilon_t^2}{\nu_t^2}\right)$$

where $\nu_t^2 = \frac{\sigma^2}{2\beta}\left(e^{\beta}-1\right) r_{t-1}^{2\gamma}$, $\varepsilon_t = r_t - \frac{\alpha}{\beta}\left(e^{\beta}-1\right) - e^{\beta} r_{t-1}$, for the discrete model of interest rates:

$$r_t = e^{\beta} r_{t-1} + \frac{\alpha}{\beta}\left(e^{\beta}-1\right) + \varepsilon_t \quad (t = 2, ...N),$$

which is derived from the stochastic equation

$$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dw;$$

is proposed in [67] by Stehlíková. There is presented the condition under which the maximum exists. If the data indicate a very high level of mean reversion, then this condition is violated.

An artificial example of the overnight time series $r_t, t = 1, 2, ...N$ for which the LF has no maximum is as follows

$$r_t = a + b\frac{(-1)^t}{t},$$

where a, b are positive constants. This example is shown on Figure 3.2.

If we thought that such an example could not exist in real data set, we are wrong. During the calibration of one factor interest rate model on BRIBOR data Stehlíková

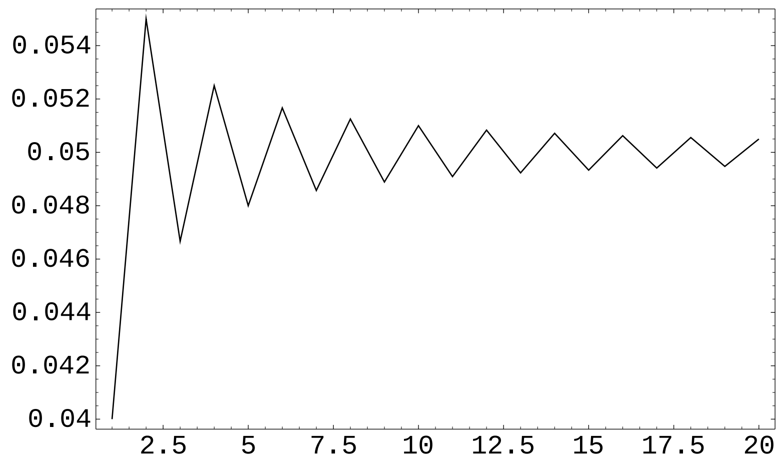


Figure 3.2: Example of non-existence of the maximum of the likelihood function. Vertical axis represents modelled interest rates whereas horizontal axis represents the time.

detected a period (December 2004) during which the condition mentioned above was not fulfilled. It means that the maximum likelihood estimates did not exist.

In the one factor term structure model the state variable is not observable, because in the market does not exist any instrument which is maturing at the next instant. So the conditional density is not known and the maximum likelihood estimator could not be used. Gibbons and Ramaswamy [37] used instead the steady-state density of the interest rate, while Brown and Dybving [17] used neither the conditional nor the steady-state density to avoid this problem. On the other hand the problem could be solved in another way, e.g. Chan, Karolyi, Longstaff and Sanders [20] used the one-month Treasury yield as a proxy for the instantaneous interest rate.

A lot of effort is being spent on the cases, when the state variable is known (or is assumed to be known) but the likelihood function is intractable. Pedersen [57] and Santa-Clara [60], Brandt and Santa-Clara [18], Durham and Gallant [30] proposed the Simulated maximum likelihood estimation (SMLE), through which we can approximate the likelihood function. The main idea of the SMLE is to split each observation interval into small subintervals and to simulate a large number of path for these subintervals. The average of the normal likelihoods in limit converges to the true likelihood of moving from one observed value to the next. Ait-Sahalia [1, 2] developed a series of approximations to the likelihood function that are tractable to estimate and converge to the true likelihood function. Brandt and He [16] developed a method of calculation on an approximate likelihood function for certain classes of term structure models.

3.3 Generalized method of moments

Generalized method of moments (GMM) by Hansen [38] is an econometric approach also used in estimating the parameters of the interest rate models. The GMM framework gives space not only for estimation, but for testing different versions of interest rate models.

This methodology uses discrete-time econometric version of the short term interest rate process. It means that in the case of the following stochastic differential equation for the short rate process:

$$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dw$$

the discrete-time specification is as follows:

$$\begin{aligned} r_{t+1} - r_t &= \alpha + \beta r_t + \epsilon_{t+1} \\ E(\epsilon_{t+1}) &= 0 \\ E((\epsilon_{t+1})^2) &= \sigma^2 (r_t)^{2\gamma} \end{aligned}$$

where $E(.)$ is the mean of the process. The unknown parameters are $\theta = (\alpha, \beta, \sigma^2, \gamma)$. This method requires the estimation of the above mentioned parameters by the equation $E[f_t(\theta)] = 0$, where

$$f_t(\theta) = (\epsilon_{t+1}, \epsilon_{t+1} r_t, \epsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}, (\epsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}) r_t)'.$$

The GMM consists of replacing $E(f_t(\theta))$ with its sample counterpart

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} f_t(\theta),$$

where T is the number of observations, and choosing parameter estimates to minimize the quadratic form:

$$J_T(\theta) = g_T^{'}(\theta) W_T(\theta) g_T(\theta)$$

for some positive definite weighting matrix $W_T(\theta)$.

This method has several advantages. Firstly, it does not require that the distribution of interest rate changes is normal and the asymptotic justification for this method requires only that the distribution of interest rate changes is stationary and ergodic and that the relevant expectations exist. This property is important because some models assume that the interest rate changes are normal, like in our case Vašiček model, but others assume non-central χ^2 variate like the CIR model. Secondly, the GMM estimator and its standard errors are consistent in the case of conditionally heteroscedastic disturbances, too.

The negative side of this approach is the impact of aggregation problem arising from the simplification by discretization, on the parameter estimates because of the influence on the distribution of disturbances.

Table 3.1: Pros and cons of different calibration and estimation techniques

methods		pros and cons
MCMC	$-$	• distribution $p(\theta, X\|Y)$ is of very high dimension • a-priori specification of the proposal density function • a-priori specification of the prior distribution (especially the non-informative priors) • term structure models (non-linearity, stochastic singularities)
	$+$	• capability to estimate multi factor models • suitable for non-normal and non-standard distribution of the variables
MLM	$-$	• required knowledge of the state variable (conditional density of the variable) • non-existence of the maximum of likelihood function in some cases
	$+$	• time-series and cross-sectional approach or the combination of these can be used • closed form of the LF is known in some cases
GMM	$-$	• aggregation problem because of discretization • do not use the whole term structure, only the overnight interest rate
	$+$	• framework for testing of models • does not require normality of interest rate changes

The GMM approach has been used by number of researchers for studying, calibrating and comparing interest rate models e.g. Gibbons and Ramaswamy [36], Harvey [39], Longstaff [50], Chan, Karolyi, Longstaff and Sanders [20], Li [49] and Olšarová [54].

Every calibration methodology has its advantages and disadvantages as it was described above. In some cases the assumptions of the applicability are very strict and can not be fulfilled in the real world or there are strong simplifications which make big errors. We summarize the problematic areas in each calibration and estimation methodology in Table 3.1.

Chapter 4

Goals of the thesis

The correct pricing of interest rate derivatives is very important in the financial business. There are many attempts spent on this topic. Several approaches have exist, as it was summarized in the previous chapter, but no definite pricing model has been proposed. Many researchers calibrate the interest rate dependent securities through their market prices. Others try to estimate the interest rate models parameters directly and then utilize this information in the pricing of the derivatives. If one chooses the first approach then he needs a market with different securities. This methodology is problematic in the new EU member states because of their less active financial markets. They are characterized by strong dominance of the banking sector. This means that we can calibrate the interest rate models through bonds term structure.

Now it is important to choose a calibration technique for this issue. Some problems, depicted in the previous chapter, with different estimation and calibration methodologies lead us to propose a new, the so called minmax optimization method. This procedure consists of two steps. In the first step we minimize the so-called loss functional. It leads to a one-dimensional λ-parameterized curve of minimizers. In the second step we maximize the likelihood function restricted to this curve. Measuring the quality of this method is also important, so we introduce two ratios the non-linear R^2 ratio and the maximum likelihood ratio.

As an extension of this method we introduce the binding of term structures. It utilizes more (statistical) informations recieved from the yield curves trying to bound the mean or the whole term structure. For that reason the results are in the form of intervals.

Of course, basic idea of calibration is the reduction of parameter space and the proposition of transformed parameters.

The comparison of the results of calibration for Central and Western European countries is one of the aims of this thesis.

Our attention is primary focused on the internal calibration method and its possible application to different data basis (which need not be necessarily up-to-date). After the internal calibration method we present the external method which basically assume a potential extra information provided externally. That means, the calibration is performed not only on the data basis (or informations acquired from the data), but also e.g. an expert judgment. In this case we obtain not a point of calibrated parameters but an interval of possible values.

Our goal is not only to present a calibration method which can be used on the one factor models, but also suggest a possible methodology for the multi factor models. The comprehensive demonstration of the proposed methods is rather difficult on the multi factor models, the understanding and interpretation of results is complex. The main reason for the selected approach is to extract the maximum of informations from the one factor models, especially the CIR and the Vašiček model, and also indicate the possible way for the multi factor models calibration in the presented framework. According to the obtained information we would be able to comment our results from the positive and also negative point of view, find out some advantages and disadvantages.

Most of the results on internal calibration methods are new and were developed and published by Ševčovič and Urbánova-Csajková just recently [64, 65].

Chapter 5

The loss functional and its minimization

In the one factor interest rate models, the only stochastic factor is the spot rate following the mean reverting process of the form (2.6). As it was derived in Section 2.3, the governing equation for the zero-coupon bond price is:

$$\frac{\partial P}{\partial t} + (\kappa(\theta - r) - \widetilde{\lambda}\sigma r^{\gamma})\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 P}{\partial r^2} - rP = 0, \tag{5.1}$$

where $t \in (0, T)$ and $r > 0$. The form of the parameter $\widetilde{\lambda}$ (the market price of risk) distinguishes the Vašiček from the CIR model. We also know that an explicit solution to the above mentioned PDE exists in the form: $P(T-\tau, T, r) = A(\tau)e^{-B(\tau)r}$, where $A(\tau)$ and $B(\tau)$ are defined in (2.14) for the Vašiček and (2.15) for the CIR model. As we can see in these models there are four unknown parameters: $(\kappa, \sigma, \theta, \lambda)$. Notice that in the CIR model, the market price of risk λ appears only in summation $\kappa + \lambda$. Key feature of our next work was presented also by Pearson and Sun in [56]. This was the first motivation to look properly at the parameters involved in the above mentioned models.

Another motivation comes from the option pricing theory. In this case we can reduce one parameter, the expected rate of return μ. It means that the Black-Scholes partial differential equation for the option price does not depend on the risk preferences of the investor. The option price exactly depends on three parameters (σ, r, E) instead of four (σ, r, E, μ), i.e. $V(S, t; \sigma, r, E)$ (see [47]).

In the bond pricing theory the underlying instrument is the short rate. This is not a tradable instrument so it could not be used in the hedging and for that the market preferences of the investors could not be fully eliminated from the bond price. But in the following section we will see that some parameter reduction can be made.

5.1 Parameter reduction

There are many ways how to reduce the original four dimensional parameters into three reduced parameters. Our objective is however to find a transformation leading to three-parameters reduction.

The idea of reducing the four dimensional parameter space into three parameters is possible in the case of CIR and Vašiček model, too.

5.1.1 Case of the Cox-Ingersoll-Ross model

The parameter reduction for the CIR model consists of introduction of the following set of new variables:

$$\beta = e^{-\eta}\,, \quad \xi = \frac{\kappa+\lambda+\eta}{2\eta}\,, \quad \varrho = \frac{2\kappa\theta}{\sigma^2}, \tag{5.2}$$

where $\eta = \sqrt{(\kappa+\lambda)^2 + 2\sigma^2}$. Returning back to the original CIR parameters $(\kappa, \sigma, \theta, \lambda)$ we have

$$\kappa = \eta(2\xi - 1) - \lambda\,, \quad \sigma = \eta\sqrt{2\xi(1-\xi)}\,, \quad \theta = \frac{\varrho\sigma^2}{2\kappa}, \tag{5.3}$$

where $\eta = -\ln\beta$. The following proposition is now a consequence of explicit formulae (2.15).

Proposition 5.1.1. *In terms of transformed parameters the value of a bond $P = P(T - \tau, T, r)$ can be expressed as $P = Ae^{-Br}$, where $\tau = T - t \in [0, T]$ and functions $A = A(\beta, \xi, \varrho, \tau),\ B = B(\beta, \xi, \varrho, \tau)$ satisfy*

$$B = -\frac{1}{\ln\beta}\,\frac{1-\beta^\tau}{\xi(1-\beta^\tau)+\beta^\tau}\,, \qquad A = \left(\frac{\beta^{(1-\xi)\tau}}{\xi(1-\beta^\tau)+\beta^\tau}\right)^{\varrho}. \tag{5.4}$$

Moreover, $(\beta, \xi, \varrho) \in \Omega = (0,1)\times(0,1)\times\mathbb{R}^+ \subset \mathbb{R}^3$.

It is convenient to introduce the transformation $T : \mathcal{D} \to \Omega$ defined as in (5.2) where $\mathcal{D} = (0,\infty)^3 \times \mathbb{R} \subset \mathbb{R}^4$. Then $T(\kappa, \sigma, \theta, \lambda) = (\beta, \xi, \varrho)$, is a smooth mapping and, for any $(\check\beta, \check\xi, \check\varrho) \in \Omega$, the preimage

$$T^{-1}(\check\beta, \check\xi, \check\varrho) = \{(\kappa_\lambda, \sigma_\lambda, \theta_\lambda, \lambda) \in \mathbb{R}^4,\ \lambda \in \check J\}, \quad \check J = (-\infty, -(2\check\xi - 1)\ln\check\beta),$$

is a smooth one-dimensional λ-parameterized curve in $\mathcal{D} \subset \mathbb{R}^4$ where

$$\begin{aligned}
\kappa_\lambda &= -\lambda - (2\check\xi - 1)\ln\check\beta, \\
\sigma_\lambda &= -\sqrt{2\check\xi(1-\check\xi)}\ln\check\beta, \qquad (5.5) \\
\theta_\lambda &= \frac{\check\varrho\sigma_\lambda^2}{2\kappa_\lambda}, \qquad (5.6)
\end{aligned}$$

where $\lambda \in \check{J}$.

5.1.2 Case of the Vašiček model

As far as the Vašiček model is considered we put

$$\beta = e^{-\kappa}\,, \quad \xi = \theta - \frac{\sigma^2}{2\kappa^2} - \frac{\sigma\lambda}{\kappa}\,, \quad \varrho = \frac{\sigma^2}{4\kappa}. \tag{5.7}$$

Then for the original Vašiček parameters we have:

$$\kappa = -\ln\beta\,, \quad \sigma = 2\sqrt{\varrho\kappa}\,, \quad \theta = \xi + \frac{\sigma^2}{2\kappa^2} + \frac{\sigma\lambda}{\kappa}. \tag{5.8}$$

Similarly as in above, it follows from (2.14):

Proposition 5.1.2. *In terms of transformed parameters the value of a bond $P = P(\tau, r)$ can be expressed as $P = Ae^{-Br}$, where $\tau = T - t \in [0, T]$ and functions $A = A(\beta, \xi, \varrho, \tau),\ B = B(\beta, \xi, \varrho, \tau)$ satisfy*

$$B = -\frac{1-\beta^\tau}{\ln\beta}\,, \quad A = \exp\left(\xi(B(\tau) - \tau) - \varrho B^2(\tau)\right)\,. \tag{5.9}$$

where $(\beta, \xi, \varrho) \in \Omega = (0,1) \times \mathbb{R} \times \mathbb{R}^+ \subset \mathbb{R}^3$.

The transformation $T : \mathcal{D} \to \Omega$ defined as in (5.7), i.e. $T(\kappa, \sigma, \theta, \lambda) = (\beta, \xi, \varrho)$, where $\mathcal{D} = (0, \infty)^3 \times \mathbb{R} \subset \mathbb{R}^4$, is a smooth mapping too and, for any $(\check\beta, \check\xi, \check\varrho) \in \Omega$, the preimage

$$T^{-1}(\check\beta, \check\xi, \check\varrho) = \{(\kappa_\lambda, \sigma_\lambda, \theta_\lambda, \lambda) \in \mathbb{R}^4,\ \lambda \in \check{J}\}, \quad \check{J} = \mathbb{R},$$

is a smooth one-dimensional λ-parameterized curve in $\mathcal{D} \subset \mathbb{R}^4$. In this case

$$\kappa_\lambda = -\ln\check\beta, \quad \sigma_\lambda = 2\sqrt{\check\varrho\kappa_\lambda}, \quad \theta_\lambda = \check\xi + \frac{\sigma_\lambda^2}{2\kappa_\lambda^2} + \frac{\sigma_\lambda\lambda}{\kappa_\lambda}\,. \tag{5.10}$$

Summarizing, in both studied one factor models the yield curve depends only on three transformed parameters β, ξ and ϱ defined in (5.2) and (5.7), resp.

5.2 Loss functional

In this section we introduce the loss functional which measures the quality of approximation of the set of real market yield curves by computed yield curves from each model.

Definition 5.2.1. *The loss functional is the time-weighted distance of the real market yield curves* $\{R_j^i, j = 1, ..., m\}$ *and the set of computed yield curves* $\{\bar{R}_j^i, j = 1, ..., m\}$ *at time* $i = 1, ..., n,$ *determined from the bond price - yield curve relationship*

$$A_j e^{-B_j R_0^i} = e^{-\bar{R}_j^i \tau_j} \tag{5.11}$$

where $r^i = R_0^i$ *is the overnight interest rate at time* $i = 1, ..., n,$ $A_j = A(\tau_j)$ *and* $B_j = B(\tau_j)$ *where* $0 = \tau_0 < \tau_1 < \tau_2 < ... < \tau_m$ *stand for maturities of bonds forming the yield curve, is defined as follows:*

$$U(\beta, \xi, \varrho) = \frac{1}{m} \sum_{j=1}^{m} \frac{1}{n} \sum_{i=1}^{n} (R_j^i - \bar{R}_j^i)^2 \tau_j^2. \tag{5.12}$$

Recall that $A(\tau)$ and $B(\tau)$ are defined by (5.4) and (5.9).

Proposition 5.2.1. *In terms of the averaged term structure values and their covariance values the loss functional can be expressed in form:*

$$\begin{aligned} U(\beta, \xi, \varrho) &= \frac{1}{m} \sum_{j=1}^{m} ((\tau_j E(R_j) - B_j E(R_0) + \ln A_j)^2 \\ &+ D(\tau_j R_j - B_j R_0)), \end{aligned} \tag{5.13}$$

where $E(X_j)$ *and* $D(X_j)$ *denote the mean value and dispersion of the vector* $X_j = \{X_j^i, i = 1, ..., n\}$.

This equivalent form of the loss functional is derived in details in the Appendix 5.7 (to Chapter 5).

Expression (5.13) for the loss functional is much more suitable for computational purposes because it contains aggregated time series information from the yield curve only, the cumulative statistics like the mean and covariance of term structure R_j series. These statistical informations can be pre-processed prior to optimization.

5.3 Non-linear regression problem for the loss functional

Introducing the short form of the loss functional (5.13) is prerequisition to the next steps. The core of the estimation method is to minimize the following function:

$$\min_{(\beta, \xi, \varrho) \in \Omega} U(\beta, \xi, \varrho)$$

where $\Omega = (0, 1) \times (0, 1) \times (0, \varrho_{max})$ [1] is a bounded domain in $\mathbb{R}^3$. During this step of our approach we obtain the vector of $(\check{\beta}, \check{\xi}, \check{\varrho})$ for any given λ. This problem is

[1] ϱ_{max} is sufficiently large; in our computation $\varrho_{max} = 5$.

highly non-linear. For that reason we discuss different numerical procedures in the next section. Having identified the curve of global minimizers of the loss functional we proceed by the second step which will be discussed later in Section 6.1 and 6.2.

Remark 5.3.1. *The parameter reduction described in the previous section can be followed by optimal selection of some of the parameters (see Propositions 5.3.1 and 5.3.2).*

5.3.1 Case of the Cox-Ingersoll-Ross model

Proposition 5.3.1. *Given β and ξ, an optimal value for the parameter ϱ in the CIR model $\varrho_c^{opt} = \varrho_c^{opt}(\beta, \xi)$ can be found as a function of β and ξ. Solving the first order optimality condition $\frac{\partial U}{\partial \varrho} = 0$ we have:*

$$\sum_{j=1}^{m} (\ln A_j)^2 = -\sum_{j=1}^{m} (\tau_j E(R_j) - B_j E(R_0)) \ln A_j \tag{5.14}$$

and the optimal ϱ_c is determined as follows:

$$\varrho_c^{opt} = -\frac{\sum_{j=1}^{m} (\tau_j E(R_j) - B_j E(R_0)) \ln A_j(\beta, \xi, 1)}{\sum_{j=1}^{m} (\ln A_j(\beta, \xi, 1))^2}. \tag{5.15}$$

The proof can be find in Appendix 5.7.

5.3.2 Case of the Vašiček model

Proposition 5.3.2. *Given β, a pair of optimal values for the parameter (ϱ, ξ) in the Vašiček model $\varrho_v^{opt} = \varrho_v^{opt}(\beta)$, $\xi_v^{opt} = \xi_v^{opt}(\beta)$ can be found. Solving the system of first order optimality conditions $\frac{\partial U}{\partial \varrho} = 0$ and $\frac{\partial U}{\partial \xi} = 0$ we have:*

$$\begin{aligned} 0 &= \sum_{j=1}^{m} (\tau_j E(R_j) - B_j E(R_0) + \xi(B_j - \tau_j) - \varrho B_j^2) B_j^2 \\ 0 &= \sum_{j=1}^{m} (\tau_j E(R_j) - B_j E(R_0) + \xi(B_j - \tau_j) - \varrho B_j^2)(B_j - \tau_j) \end{aligned} \tag{5.16}$$

and the pair of optimal values $(\varrho_v^{opt}, \xi_v^{opt})$ can be determined from the system of linear equations:

$$\begin{aligned} \varrho_v^{opt} &= \frac{\sum_{j=1}^{m} (\tau_j E(R_j) - B_j E(R_0) + \xi_v^{opt}(B_j - \tau_j)) B_j^2}{\sum_{j=1}^{m} B_j^4} \\ \xi_v^{opt} &= -\frac{\sum_{j=1}^{m} (\tau_j E(R_j) - B_j E(R_0) - \varrho_v^{opt} B_j^2)(B_j - \tau_j)}{\sum_{j=1}^{m} (B_j - \tau_j)^2}. \end{aligned} \tag{5.17}$$

The optimal values of ϱ_c^{opt} for the CIR model and $(\varrho_v^{opt}, \xi_v^{opt})$ for the Vašiček model can be picked and used during the optimization only when they are positive. This sufficient assumption is derived in Theorem 5.7.1.

Summarizing, for the CIR as well as for the Vašiček model we have first order necessary conditions for the minimizer of the loss functional. These conditions can be used either for further parameter reduction of the problem (2D problem for the CIR model and even 1D problem for the Vašiček model) or for testing whether a numerical approximation is close to a minimizer. Latter property has been used in practical implementation of the minimization method.

5.4 Numerical procedure for minimization of the loss functional

In this section we discuss several optimization methods for finding the minimum of the loss functional U on Ω. The objective is to find a numerical approximation of the optimization problem:

$$\min_{x\in\Omega} U(x) \tag{5.18}$$

where x is the vector of the unknown parameters and $\Omega \in \mathbb{R}^n$. In our case $n = 3, x = (\beta, \xi, \varrho)$ and $\Omega = (0,1) \times (0,1) \times (0, \varrho_{max})$ is a bounded domain in $\mathbb{R}^3$ where $\varrho_{max} > 0$ is sufficiently large number.

In the following we discuss three different methods we have used in order to find a minimum of the loss functional $U = U(\beta, \xi, \varrho)$.

- a steepest descend method of Newton-Kantorovich type,
- an evolution strategy based method,
- combination of these two methods.

5.4.1 A steepest descend method of Newton-Kantorovich type

This method is often used in convex optimization problems. The basic idea of this method is read as follows:

1. we choose an initial approximation x_0 of arg $\min_\Omega U$
2. we construct a sequence of approximations:

$$x_{i+1} = x_i - [\nabla^2 U(x_i)]^{-1} \nabla U(x_i), \quad i = 0, 1, 2, ... \tag{5.19}$$

where $\nabla U(x_i)$ is the gradient of U in the i-th approximation of x

3. we continue computation of x_i until a prescribed accuracy goal measured by the norm of $\|\nabla U(x_i)\|$ is attained.

The numerical approximation of partial derivatives $\partial_{x_k} U$ can be approximated by central finite differences, i.e. $\partial_{x_k} U(x) \approx (U(x+he_k)-U(x-he_k))/2h$ where $0 < h \ll 1$ is sufficiently small and $e_k \in R^n$ is the k-th vector of the canonical orthonormal basis in R^n. Second derivatives appearing in $\nabla^2 U$ are approximated similarly by means of second central differences.

According to [3] the above algorithm constructs a sequence $\{x_i\}$ converging to a local minimum of U provided the function U satisfies suitable regularity conditions and the initial condition x_0 is chosen appropriately. In this case the convergence is locally quadratical. However, in general (if the function U is not convex) the above method need not guarantee convergence and it may converge to a local minimum only (see Algower & Georg [3] for details).

Let us mention several difficulties we had to overcome when implementing this method:

1. The loss function $U(\beta, \xi, \varrho), \quad (\beta, \xi, \varrho) \in \Omega \in \mathbb{R}^3$ is not necessarily convex. The plot of this function indicate that it can be flat in one parameter. The demonstration of this feature is shown on Figure 5.1. As we can see, the flat shape is typical for the loss functional. This behaviour makes problems during the minimization not only for the Newton-Kantorovich type method but also for the evolution strategies.

2. The admissible values of optimal parameters have to satisfy the following conditions

   - $\beta \in (0, 1)$, $\varrho \in (0, \varrho_{max})$ and $\xi \in \mathbb{R}$ for the Vašiček, model
   - $\beta \in (0, 1)$, $\varrho \in (0, \varrho_{max})$ and $\xi \in (0, 1)$ for the CIR model.

   The sequence constructed as in (5.19) does not necessarily satisfy these bounds and therefore it may converge to a minimum which is not admissible from financial point of view.

The reason of this setting is the required positivity of $B(\tau)$ in $P(T-\tau) = A(\tau)e^{-B(\tau)r}$ and requirement that $\kappa, \theta, \sigma > 0$ for original variables. If $B(\tau) < 0$ then the price of the bond is increasing with increasing interest rate which is inconsistent with the bond pricing theory.

In order to achieve the required bounds $\beta, \xi \in (0, 1)$ we employ the 1-periodic extension mapping of the identity function. To ensure positivity of ϱ we use a smooth approximation of the absolute value function. Indeed, let us define a smooth approximation g of the step function $\xi - [\xi]$, i.e

$$g(\xi) \approx \xi - [\xi] \text{ for } \xi \notin N\,, \quad \text{and we put} \quad |\varrho|_h = \sqrt{\varrho^2 + h^2} - h$$

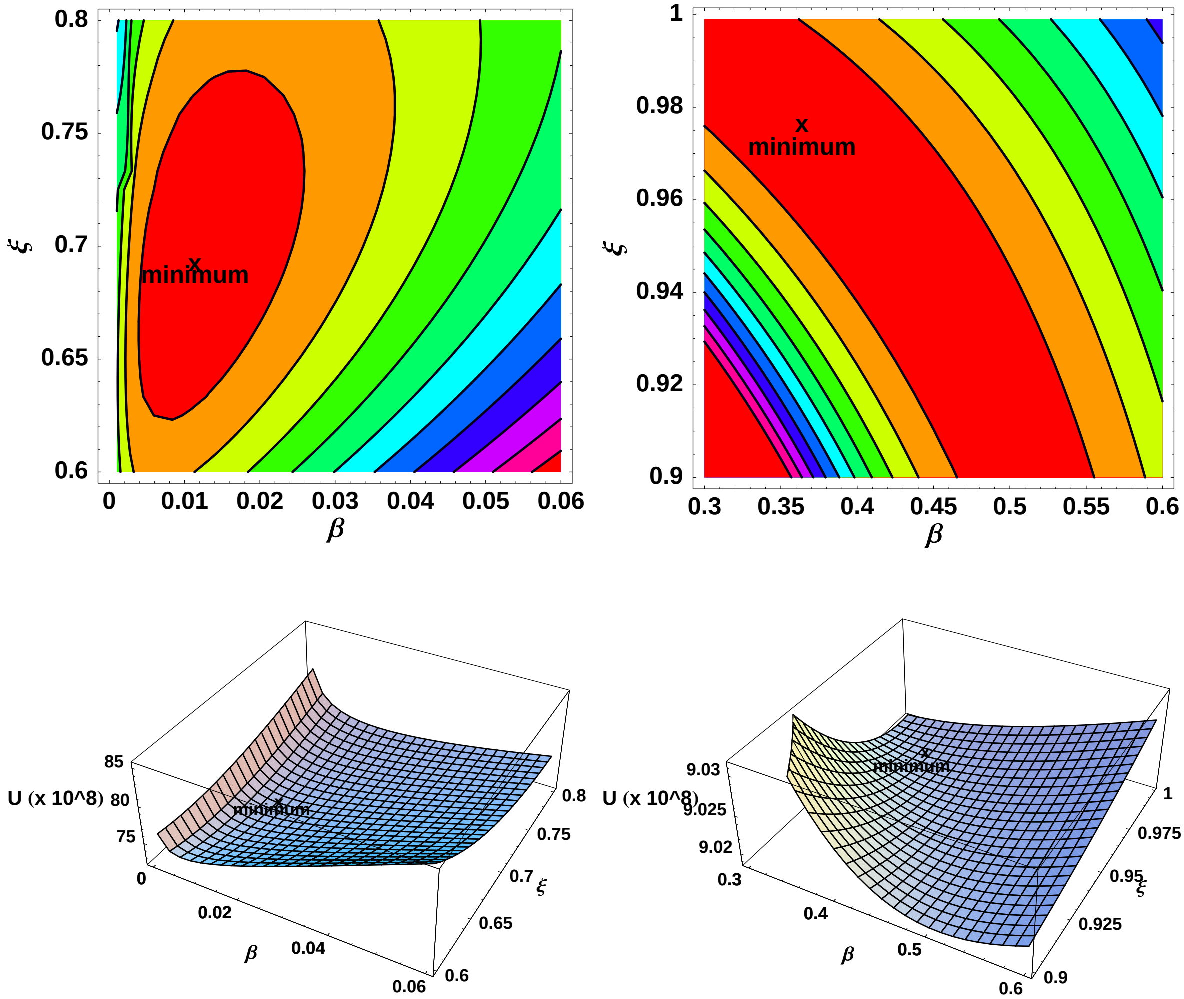


Figure 5.1: 3D plot (values have been multiplied by 10^8 so correspond to 10^{-8}) and 2D density plot of $f(\beta,\xi) = U(\beta,\xi,\varrho_{opt}(\beta,\xi))$ for Bribor 2/2005 (first column) and Pribor 7/2005 (second column).

where $[.]$ denotes the integer part, $0 < h \ll 1$ is a small parameter. Clearly, $g : R \mapsto [0,1)$ and $g(\xi) \simeq \xi$ for $\xi \in [0,1)$. Then it should be obvious that any local minimum $(\check{\beta}, \check{\xi}, \check{\varrho}) \in R^3$ of the extended minimization problem defined on the whole domain R^3

$$\min_{(\beta,\xi,\varrho)\in R^3} U(g(\beta), g(\xi), |\varrho|_h)$$

for the CIR model and

$$\min_{(\beta,\xi,\varrho)\in R^3} U(g(\beta), \xi, |\varrho|_h)$$

for the Vašiček model, resp., corresponds to a local minimum $(g(\check{\beta}), g(\check{\xi}), |\check{\varrho}|_h) \in (0,1) \times (0,1) \times (0,\infty)$ for the CIR model, and $(g(\check{\beta}), \check{\xi}, |\check{\varrho}|_h) \in (0,1) \times R \times (0,\infty)$ for the Vašicek model.

The steepest descent method of Newton-Kantorovich type has two problematic features. The first is the choice of the starting point $x_0 = (\beta_0, \xi_0, \varrho_0)$ of the computation and the second is the number of steps to achieve the predetermined accuracy. For that reason we have considered a new method which is a combination of an evolution strategy method and the steepest descend method. It will be discussed later.

5.4.2 Evolution strategies

As it was mentioned in previous section a steepest-descent methods of Newton-Kantorovich type (cf. [3]) may capture a local minimum only. This is why we have to consider a different robust numerical method generically converging to a global minimum of U. There is a wide range of optimization methods based on stochastic optimization algorithms.

These methods are often referred to as Evolution strategies (ES) (see e.g. [58, 61, 62, 63]). The main concept of this strategy is based on the survival of the fittest. There exist many different types of this stochastic algorithm like the two membered $(1+1)$ ES, multi-membered (p,c) ES, $(p+c)$ ES (see [58, 61, 62, 63]).

In our case we used a slight modification of the well known $(p+c)$ ES [58]. Recall that the $(p+c)$ ES has p parents and c children (offsprings) per population among which the p best individuals are selected to be next generation parents by their fitness value. The procedure is repeated until some termination criterion is satisfied.

The mathematical description of the modification of $(p+c)$ ES called $(p+c+d)$ ES is as follows:

The problem is defined as finding the real valued vector $x \in \Omega$ which is a global minimum of objective function U in $\Omega \subset R^n$.

1. The initial population of parent vectors $x_k \in \Omega$, $k = 1, ..., p$ is generated randomly from bounded three dimensional space $\Omega_b = \{(\beta, \xi, \varrho) \in \Omega, 0 \le \varrho \le \varrho_{max}\}$ where ϱ_{max} is large enough. Ω_b is a subset of the domain Ω.

2. In each step of the ES algorithm we generate a set of c offsprings from the parent population ($c \leq p$). Each vector of children (offspring) $\bar{x}_l$, $l = 1, ..., c$ is created from parents x_k, $k = 1, ..., p$ by mutation and recombination. Mutation means perturbation of parent generation x_k, $k = 1, ..., p$ by Gaussian noise with zero mean and preselected standard deviation σ_{gauss}. Recombination means crossing over parts of randomly chosen vectors of children.

3. The modification $(p + c + d)$ ES comprise selection on a wider set. It means that we include a randomly generated set of d wild type individuals forming the so-called wild population. The procedure of generation of the wild type population x_o, $o = 1, ..., d$, from bounded space Ω_b is the same as for the initial population.

4. Every member of the population (parents, children, wild population) is characterized with its fitness value which is the value of the loss functional U.

5. Selection chooses p best vectors from the population by their fitness value to be next generation parents. A set of p intermediate parents is obtained.

6. Next we include a corrector step consisting of improving the set of p intermediate parents by NK iterates of the Newton-Kantorovich gradient minimization method. As a result we obtain a set of p improved parents.

7. The best p individuals from the set of p parents, p improved parents, c offsprings and d wild type individuals are selected to be the next generation of parents.

8. We repeat this procedure until the overall number of steps is less than N. We also perform the first order necessity test as described in Chapter 5.

In our computations we have chosen $N = 300$, $p = c = d = 10^5$, $NK = 30$ and $\sigma_{gauss} = 0.01$. We have not update the standard deviation according to Rechenberger's rule (see [58]) as it turned to be ineffective.

Similarly as in the case of gradient optimization methods, for a general minimized function, an ES based stochastic algorithm need not necessarily converge to a global minimum. Additional assumptions like e.g. convexity made on a minimized function are required. We are unable to verify these conditions in our particular case. Nevertheless, our numerical experience based on repeated experiments with different numerical constants indicates that the ES algorithm described above indeed converges to a global minimum of the loss functional U. Moreover, an important question concerning existence and uniqueness of a global minimum of U on Ω arises. Some theoretical considerations about the global minimum of U on Ω are in Lemma 5.7.1 and Proposition 5.7.1. Notice that data vectors $R_j, j = 0, ..., m,$ enter expression for U in terms of their means and covariances. Now if a global minimum of U is attained at several minimizers then one can perturb input data vectors R_j slightly in order to perturb their means and covariances destroying their multiplicity and achieving thus a unique global minimizer of U. Therefore, for generic data vectors $R_j, j = 0, ..., m$, it reasonable to assume that there exists a unique global minimizer of U. A rigorous proof of this feature is however not included in this thesis.

5.4.3 Other methods for solving non-linear regression problems

The non-linear optimization problem could be solved also in another way. Let us remind that we are minimizing the loss functional $U(x)$, where $x = (\beta, \xi, \varrho) \in \mathbb{R}^3$ on a bounded domain. Instead of this problem we can solve an approximation of it in the following form:

$$\min_{x \in \Omega} \tilde{U}(x)$$

where $\tilde{U}(x)$ is the linear approximation of $U(x)$ in x. It means that if $U(x)$ is defined in the following general form:

$$U(x) = \sum_i (\bar{f}(x, y_i) - f_i)^2$$

where x is the vector of parameters, y is the vector of inputs, $f \in \mathbb{R}^k$ the vector of observed outputs to the optimization problem and $\bar{f} : \Omega \times \mathbb{R}^k$ is a smooth function, then one can approximate this problem by taking the first order Taylor expansion $\bar{f}(\tilde{x}, y) + \bar{f}'_x(\tilde{x}, y)(x - \tilde{x})$ of the non-linear function $\bar{f} = \bar{f}(x, y)$ in x variable at the point $\tilde{x}$, i.e.

$$\tilde{U}(x) = \sum_i (\bar{f}(\tilde{x}, y_i) + \bar{f}'_x(\tilde{x}, y_i)(x - \tilde{x}) - f_i)^2$$

where $\tilde{x} = x^n$ is the previous approximation of the argument of the minimum of U and $\bar{f}'_x$ is the gradient of $\bar{f}$ with respect to the x variable. Now this problem is a linear regression problem in x variable that can be solved easily for a new approximation $x = x^{n+1}$ of the argument of minimum of $\tilde{U}$. This way one can obtain a sequence $\{x^n\}_{n=1}^{\infty}$ of approximations of the argument of minima of U which can be shown to be convergent under suitable assumptions made on non-linearity f. More details on this kind of linearization can be found e.g. in Pázman [55], Štulajter [72], Hornišová [41].

5.5 A'posteriori analysis of residuals. Justification of the form of the loss functional

We have defined the loss functional in Section 5.2 as a measurement of the quality of approximation of the real data by the computed one from the models (Vašiček and CIR models). As it was mentioned, the optimization problem of finding the minimum of the loss functional U on Ω is highly non-linear, despite of the fact it mimics least squares approach in linear regression method. Solving the problem (5.18) we obtain

(β, ξ, ϱ). Since computed term structures $\bar{R}_i^j = \bar{R}_i^j(\check{\beta}, \check{\xi}, \check{\varrho})$ depend only on (β, ξ, ϱ) the knowledge of the global minimizer $(\check{\beta}, \check{\xi}, \check{\varrho})$ of U is enough information for the next a'posteriori analysis of residuals.

One of the basic assumption of the simple linear regression is the normality of the distribution of residuals. Since we have constructed our loss functional as a weighted sum of squares of residuals one should justify this form by an a'posteriori analysis of the distribution of residuals.

Departures from the normality assumption can be identified in several ways like graphical checking by the histogram of residuals, analyzing of the skewness and the kurtosis of the distribution of residuals, Jarque-Bera test and other methods.

For readers' convenience we recall some basic testing method for normality of distributions:

- Using the histogram to test for normality is based on the comparison of the normal curve and the distribution of residuals. This method is fundamental and gives us basic idea, whether the distribution of residuals is normal or not.

- Skewness is the measure of the symmetry of a distribution. The sample skewness can be calculated from the central moments of a distribution as follows:

$$s = \frac{m_3}{m_2\sqrt{m_2}}$$

  where $m_r = \frac{1}{N}\sum(X - \bar{X})^r$ is the *r*-th moment about the mean $\bar{X}$, and N stands for the number of observations. If a distribution is symmetrical, then $m_3 = 0$ and $s = 0$. If a distribution is right skewed, then $s > 0$. If a distribution is left skewed, then $s < 0$.

- Kurtosis is a measure of the thickness of the tails of a distribution. It tells us something about the "peakness" of a distribution and the sample kurtosis can be calculated as follows:

$$k = \frac{m_4}{m_2^2}.$$

  The moments are defined as it was depicted above. If a distribution is normal ($\frac{m_4}{m_2^2} = 3$), then $k - 3 = 0$. If a distribution is leptokurtic (fat tails), then $k - 3 > 0$. If a distribution is platykurtic (thin tails), then $k - 3 < 0$.

- The Jarque-Bera test is used for detecting departures from normality, too. It is based on the Wald test, and is computed as follows:

$$JB = N\left[\frac{s^2}{6} + \frac{(k-3)^2}{24}\right].$$

  If $JB \geq 5.99$ the null hypothesis (that $s = 0$ and $k = 3$) is rejected at significance level 5%.

Table 5.1: Testing for normality of residuals for EURO-LIBOR interest rate

EURO-LIBOR	1/4 2003	2/4 2003	3/4 2003	4/4 2003
mean ($\times 10^{-5}$)	14.0071	0.635416	-6.17419	-19.3512
variance ($\times 10^{-7}$)	6.81257	4.38528	1.81024	1.89934
skewness	-0.336927	-0.122863	0.133453	-0.296781
kurtosis	2.95249	5.23716	5.54494	5.75616
Jarque-Bera	15.5724	172.852	234.089	180.834

Table 5.2: Testing for normality of residuals for EURIBOR interest rate

EURIBOR	1/4 2003	2/4 2003	3/4 2003	4/4 2003
mean ($\times 10^{-5}$)	2.47656	0.21192	2.68519	-1.78745
variance ($\times 10^{-7}$)	6.08292	4.01117	2.12413	2.17348
skewness	-0.463778	0.0264481	0.672902	1.13337
kurtosis	3.53435	5.75725	6.07218	7.21584
Jarque-Bera	45.1194	294.703	464.041	916.456

In our case we would like to known whether the residuals defined as:

$$\begin{aligned} e_i^j &= \tau_j(R_j^i - \bar{R}_j^i) \\ &= (\tau_j R_j^i - B_j R_0^i - \ln A_j) \end{aligned} \tag{5.20}$$

where $j = 1, \ldots, m$ and $i = 1, \ldots, n$ are normally distributed. Vector of residuals E corresponds to X in our consideration of the skewness and kurtosis and $N = m.n$. Demonstration of the testing for normality has been done in program Mathematica see Chapter 10 (content of CD ROM). We have computed the residuals for different inter-bank rates quarterly because the calibration have been done on quarterly basis (see [65]) too. The data are from the year 2003, but the calculation steps would be the same for other data (e.g. for different years).

Results of our analysis are summarized in Tables 5.1-5.4. The graphical representation can be seen on Figures 5.2-5.5. As we can see the normality test results are not satisfactory. The distribution of residuals is left skewed in many cases and also leptocurtic which means that the hypothesis of normality is rejected by the Jarque-Bera test. This fact has been neglected or disregarded in many papers, when the loss functional is of the same or very similar form as we have defined. One of the results (Table 5.3, 2nd quarter) verifies that, if the significance level is decreased to 1%, the null hypothesis (that the residuals are normally distributed) can not be rejected. We compare the value from the Jarque-Bera test with 9.21. Although in this case the type 1 error (rejecting of the null hypothesis when it is true) has decreased, but the type 2 error (not rejecting of the null hypothesis when the alternative hypothesis is true) has increased.

For that reason, the definition of the loss functional [2] is the key element of the calibration and we should not forget about the residuals. Possible solution to this

[2] Our results indicate that the least square estimation for the loss functional could be better in the

Table 5.3: Testing for normality of residuals for BRIBOR interest rate

BRIBOR	1/4 2003	2/4 2003	3/4 2003	4/4 2003
mean $(\times 10^{-4})$	9.19593	9.95761	4.28008	1.87488
variance $(\times 10^{-7})$	17.6885	20.3381	5.46829	5.17393
skewness	-0.702378	-0.120432	-0.888203	-1.14491
kurtosis	5.9977	3.58279	4.8819	4.59112
Jarque-Bera	198.185	7.075	123.061	140.596

Table 5.4: Testing for normality of residuals for PRIBOR interest rate

PRIBOR	1/4 2003	2/4 2003	3/4 2003	4/4 2003
mean $(\times 10^{-4})$	0.161299	-1.82956	-0.495298	-2.40967
variance $(\times 10^{-7})$	1.17819	2.2319	0.292456	1.08391
skewness	0.924772	-2.41783	-1.25976	-0.961701
kurtosis	10.055	8.46094	5.9829	3.97165
Jarque-Bera	480.965	481.066	146.74	41.9857

problem could be in the selection of right weights in the definition of the loss functional. On the other hand, the problem of the right weight selection and calibration with subsequent testing for normality of residuals is rather computationally difficult.

case of $U(\beta, \xi, \varrho) = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n}\left|(R_j^i - \bar{R}_j^i)\tau_j\right|^p$, where $1 < p < \infty$ and $p \neq 2$.

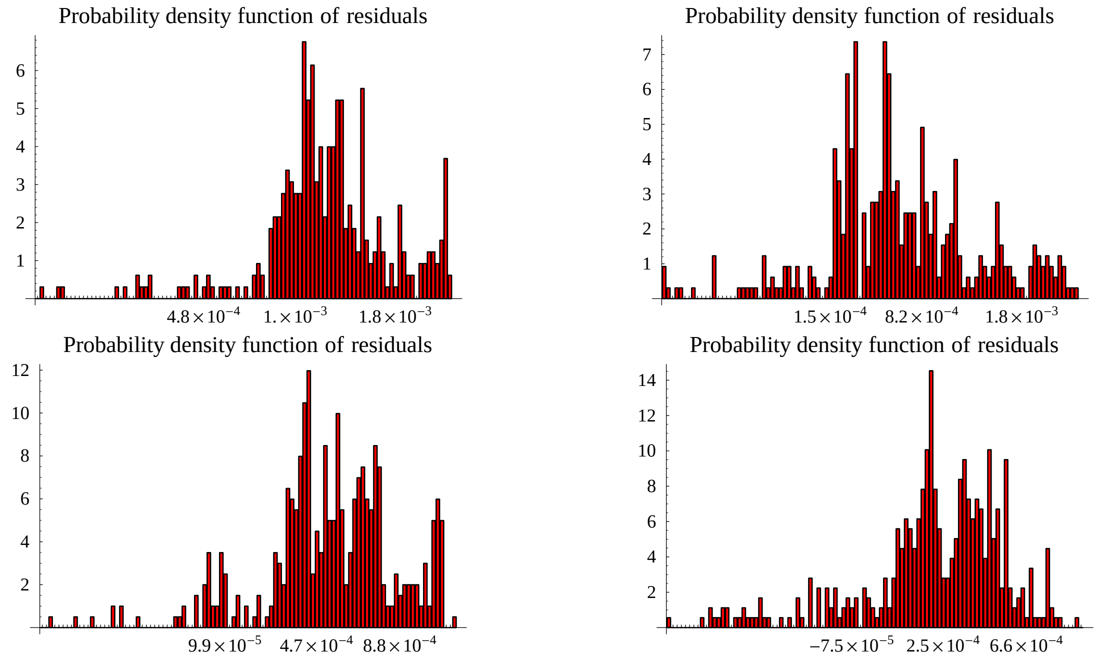


Figure 5.2: Plot of the histogram of residuals for BRIBOR rates (1Q, 2Q, 3Q, 4Q) in 2003

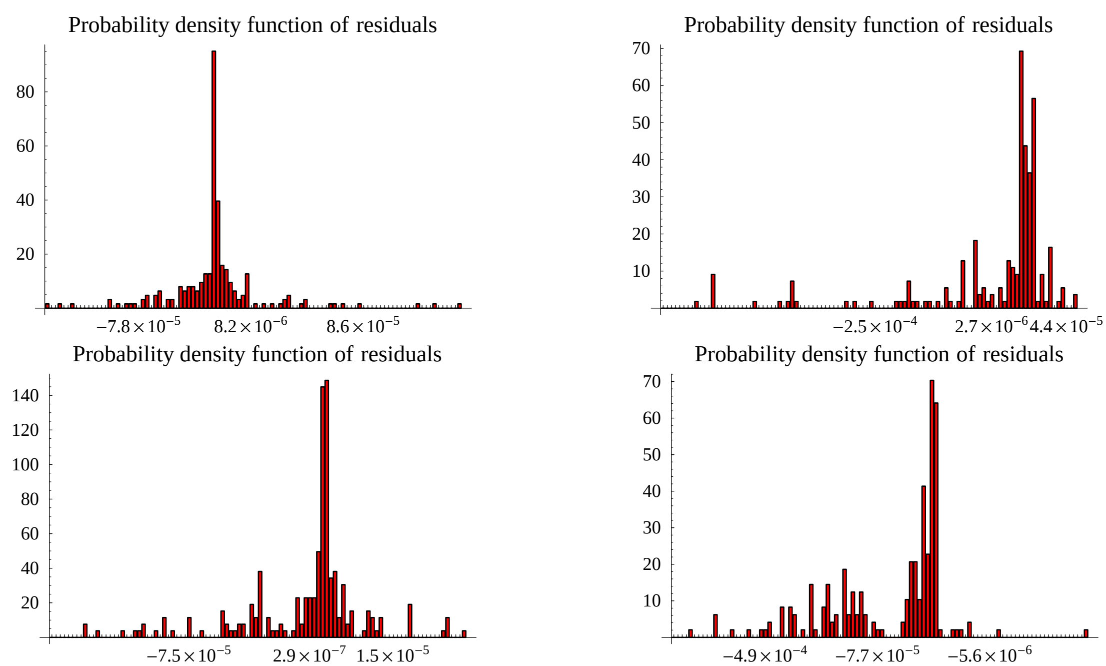


Figure 5.3: Plot of the histogram of residuals for PRIBOR rates (1Q, 2Q, 3Q, 4Q) in 2003

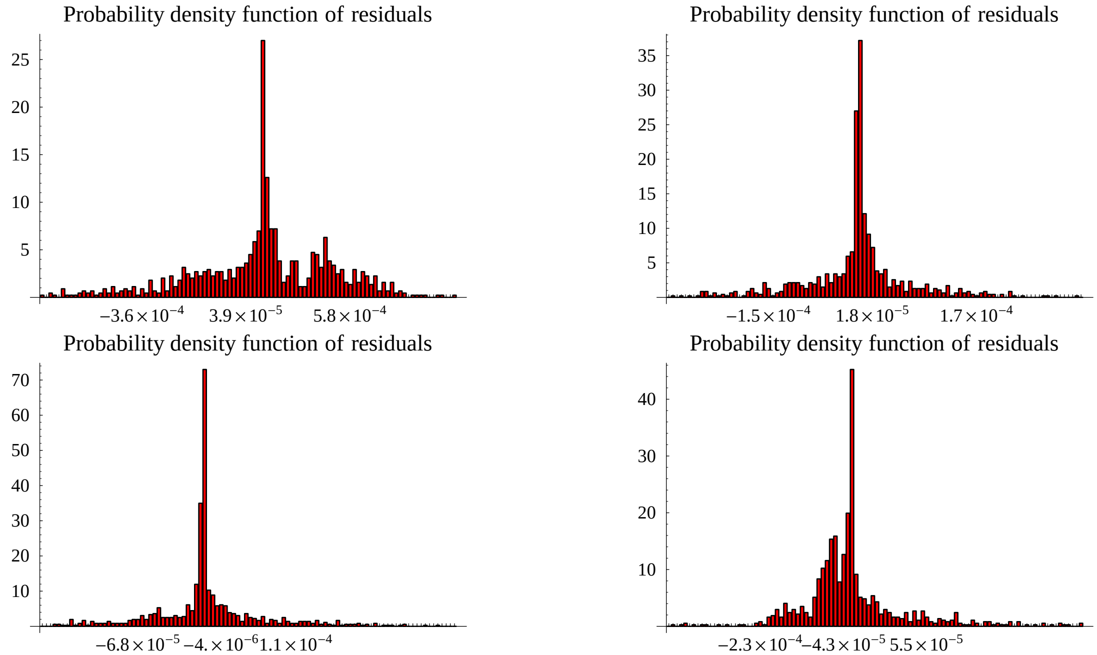


Figure 5.4: Plot of the histogram of residuals for EURIBOR rates (1Q, 2Q, 3Q, 4Q) in 2003

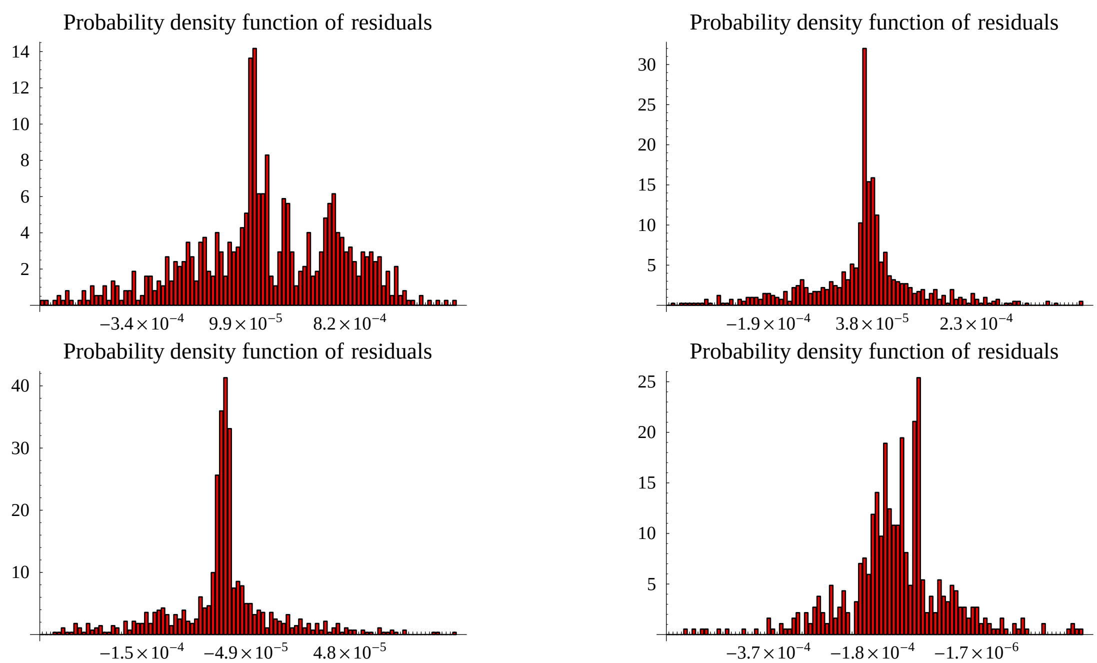


Figure 5.5: Plot of the histogram of residuals for EURO-LIBOR rates (1Q, 2Q, 3Q, 4Q) in 2003

5.6 Generalizations

As it was mentioned in Section 2.4 the solution to the two factor Fong-Vašiček model is of the following form:

$$P(\tau, r, y) = A(\tau)e^{-B(\tau)r - C(\tau)y},$$

where $A(\tau),\ B(\tau),\ C(\tau)$ satisfy the system (2.20).

Definition 5.6.1. *The loss functional, for the two factor model case which is the time-weighted distance of the real market yield curves $\{R^i_j, j = 1, ..., m\}$ and the set of computed yield curves $\{\bar{R}^i_j, j = 1, ..., m\}$ at time $i = 1, ..., n$, determined from the bond price - yield curve relationship*

$$A_j e^{-B_j R^i_0 - C_j y} = e^{-\bar{R}^i_j \tau_j} \tag{5.21}$$

where $r^i = R^i_0$ is the overnight interest rate at time $i = 1, ..., n$, $A_j = A(\tau_j),\ B_j = B(\tau_j)$ and $C_j = C(\tau_j)$ where $0 \le \tau_0 < \tau_1 < \tau_2 < ... < \tau_m$ stand for maturities of bonds forming the yield curve, is defined as follows:

$$U^{2f}(\kappa_1, \theta_1, \sigma_1, \lambda_1, \kappa_2, \lambda_2, \upsilon, \varrho) = \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n}(R^i_j - \bar{R}^i_j)^2\tau_j^2. \tag{5.22}$$

In the case of $A(\tau) = \tilde{A}(\tau)e^{-F(\tau)}$, where $\tilde{A}(\tau)$ and $F(\tau)$ satisfy the system:

$$\begin{aligned} \dot{\tilde{A}} &= -\tilde{A}\left[(\lambda_1\sigma_1 - \kappa_1\theta_1) + \frac{1}{2}\theta_2 B^2\right] && (5.23)\\ \dot{F} &= \lambda_1\sigma_1 B + \kappa_2\theta_2 C + \frac{1}{2}\theta_2 B^2 && (5.24) \end{aligned}$$

we can rewrite the solution to the Fong-Vašiček model in the following form:

$$P(\tau, r, y) = \tilde{A}(\tau)e^{-B(\tau)r - C(\tau)y - F(\tau)}$$

where the corresponding equations to the functions $B(\tau)$ and $C(\tau)$ remain the same as in (2.20).

In this case the functions $\tilde{A}(\tau)$ and $B(\tau)$ are the same as for the one factor version of the Vašiček model, so if we consider a set of parameters $\Theta = (\kappa_1, \theta_1, \sigma_1, \lambda_1, \kappa_2, \lambda_2, \upsilon, \varrho)$ such that $C(\tau) \equiv 0$ and $F(\tau) \equiv 0$ we get the one factor Vašiček model.

Proposition 5.6.1. *In terms of the one factor version of the Vašiček model the loss functional in the two factor case can be expressed in form:*

$$\begin{aligned} U^{2f}(\Theta) &= U^{v} - 2\frac{1}{m}\sum_{j=1}^{m}(C_j y + F_j)\left[\tau_j E(R_j) - B_j E(R_0) + \ln A_j\right] \\ &+ \frac{1}{m}\sum_{j=1}^{m}(C_j y + F_j)^2, && (5.25) \end{aligned}$$

where $E(X_j)$ and $D(X_j)$ denote the mean value and dispersion of the vector $X_j = \{X_j^i, i = 1, ..., n\}$ and U^v stands for the loss functional for the Vašiček model in the one factor case.

The second stochastic factor in this model is the volatility of the short rate. In practice we do not have any information about this parameter. The short rate can be approximated by the overnight interest rate, but we have no proxy of the volatility. One possible solution to this problem is to compute the average values for the model through the second stochastic value.

The averaging leads to the average loss functional $< U^{2f}(\Theta) >_y$ of the form:

$$\begin{aligned} < U^{2f}(\Theta) >_y &= U^v \\ &- 2\frac{1}{m}\sum_{j=1}^{m}(C_j < y > + F_j)\left[\tau_j E(R_j) - B_j E(R_0) + \ln A_j\right] \\ &+ \frac{1}{m}\sum_{j=1}^{m}(C_j^2 < y^2 > + 2C_jF_j < y > + F_j^2), \end{aligned} \tag{5.26}$$

The process of the short rates' volatility (2.17) has Gamma distribution in limit with the mean $\mu = \theta_2$ and the variance $var = \frac{v^2\theta_2}{2\kappa_2}$. It means that $< y >= \theta_2$ and

$$< y^2 >= var + < y >^2 = \frac{\theta_2(v^2 + 2\theta_2\kappa_2)}{2\kappa_2}$$

where $< y^m >= \int_{-\infty}^{+\infty} y^m g(y)dy$, where g is the density function of the Gamma distribution (for more details see [68]).

The calibration of the two factor model can be based on the calibration of U^{2f} to real market data. One of the possible approaches is presented in [4], where the author took the calibrated parameters of the one factor Vašiček model as an input to the two factor model calibration and was seeking for the remaining unknown variables.

5.7 Appendix to Chapter 5

Proof. (Proof of Proposition 5.2.1.) The logarithm of (5.11) is:

$$\ln A_j - B_j R_0^i = -\bar{R}_j^i \tau_j. \tag{5.27}$$

Substituting this expression to the loss functional we obtain:

$$\begin{aligned} U(\beta,\xi,\varrho) &= \frac{1}{m}\sum_{j=1}^{m}\frac{1}{n}\sum_{i=1}^{n}(R_j^i\tau_j - B_jR_0^i + \ln A_j)^2 \\ &= \frac{1}{m}\sum_{j=1}^{m}[\tau_j^2 E(R_jR_j) + B_j^2 E(R_0R_0) + (\ln A_j)^2 \\ &\quad -2\tau_j B_j E(R_jR_0) + 2\tau_j \ln A_j E(R_j) - 2B_j \ln A_j E(R_0)] \\ &= \frac{1}{m}\sum_{j=1}^{m}[(\tau_j(E(R_j)) - B_j(E(R_0)) + (\ln A_j))^2 \\ &\quad +cov(\tau_jR_j,\tau_jR_j) + cov(B_jR_0, B_jR_0) - 2cov(\tau_jR_j, B_jR_0)]. \end{aligned}$$

This imply the form defined in (5.13). □

Lemma 5.7.1. *Let* $\breve{\varrho}(\beta,\xi) = \max(0, \breve{\varrho}^{aux}(\beta,\xi))$ *then* $\forall(\beta,\xi,\rho) \in \Omega = [0,1]\times[0,1]\times[0,\infty)$ *it holds that* $U(\beta,\xi,\varrho) \geq U(\beta,\xi,\breve{\varrho}(\beta,\xi))$.

Proof. If $(\beta,\xi,\varrho) \in \Omega_L = [0,1]\times[0,1]\times\mathbb{R}$ we define

$$\breve{\varrho}^{aux} = -\frac{\sum_{j=1}^{m}(\tau_jE(R_j) - B_jE(R_0))\ln A_j(\beta,\xi,1)}{\sum_{j=1}^{m}(\ln A_j(\beta,\xi,1))^2}.$$

Let U has its global minimum on Ω_L in $(\breve{\beta},\breve{\xi},\breve{\varrho})$.

I. If $\breve{\varrho} > 0 \Longrightarrow (\breve{\beta},\breve{\xi},\breve{\varrho})$ is the global minimizer of U on $\Omega = [0,1]\times[0,1]\times[0,\infty)$.

II. If $\breve{\varrho} < 0 \Longrightarrow \forall\beta,\xi \in [0,1]$ and $\forall\varrho \in \mathbb{R}$ it holds that $U(\beta,\xi,\varrho) \geq U(\beta,\xi,\breve{\varrho}^{aux}(\beta,\xi))$.

$\forall\beta,\xi \in [0,1]$ and $\forall\varrho \geq 0$ we could distinguish two cases:

1. $\breve{\varrho}^{aux} > 0 \Longrightarrow U(\beta,\xi,\varrho) \geq U(\beta,\xi,\breve{\varrho}^{aux}(\beta,\xi))$.
2. $\breve{\varrho}^{aux} < 0 \Longrightarrow U(\beta,\xi,\varrho) \geq U(\beta,\xi,0)$.

It means that $U(\beta,\xi,\varrho) \geq U(\beta,\xi,\breve{\varrho}(\beta,\xi))$.

□

Proposition 5.7.1. *If* $(\breve{\beta},\breve{\xi},\breve{\varrho}) \in \Omega = [0,1]\times[0,1]\times[0,\infty)$ *is the global minimum of* U *on* Ω *then* $\breve{\varrho} = \breve{\varrho}(\breve{\beta},\breve{\xi})$.

Proof. For $\forall(\beta,\xi,\varrho) \in \Omega$ it holds that $U(\beta,\xi,\varrho) \geq U(\breve{\beta},\breve{\xi},\breve{\varrho})$ especially more $U(\breve{\beta},\breve{\xi},\varrho) \geq U(\breve{\beta},\breve{\xi},\breve{\varrho})$.

We have two cases for the quadratic function $\varrho \longmapsto U(\breve{\beta},\breve{\xi},\breve{\varrho})$ in parameter ϱ:

I. if $\breve{\varrho} > 0 \Longrightarrow \breve{\varrho} = \breve{\varrho}^{aux}(\breve{\beta},\breve{\xi}) = \breve{\varrho}(\breve{\beta},\breve{\xi})$.

II. if $\breve{\varrho} = 0 \Longrightarrow \breve{\varrho} = \max(0,\breve{\varrho}^{aux}) = \breve{\varrho}(\breve{\beta},\breve{\xi}) = 0$. □

Theorem 5.7.1. *(For CIR model) If $\xi \geq \frac{1}{2}$ and $E(R_j) \geq E(R_0)$ (yield curve monotonicity assumption) then $\breve{\varrho}^{aux} \geq 0$.*

Proof. If $\xi = \frac{\kappa+\lambda+\eta}{2\eta} \geq \frac{1}{2} \Longleftrightarrow \kappa + \lambda \geq 0 \Longleftrightarrow \lambda \geq -\kappa$ which is the limited market price of risk, then $B(\tau) \leq \tau$.

Let us show this implication:

$$\begin{aligned} -\frac{1}{\ln\beta}\frac{1-\beta^\tau}{\xi(1-\beta^\tau)+\beta^\tau} &\leq \tau \\ \frac{1-\beta^\tau}{\xi(1-\beta^\tau)+\beta^\tau} &\leq -\tau\ln\beta \end{aligned} \tag{5.28}$$

We use the following substitution: $x = \beta^\tau$:

$$(1-x) \leq -[\xi(1-x)+x]\ln x \tag{5.29}$$

Denote $F(x) = (1-x) + [\xi(1-x)+x]\ln x$. We require $F(x) \leq 0$. If $x = 1$ then $F(1) = 0$.

$$\frac{d}{dx}F(x) = \xi(\frac{1}{x}-1) + (1-\xi)\ln x$$

$$\frac{d^2}{dx^2}F(x) = -\frac{\xi}{x^2} + \frac{1-\xi}{x}$$

We require the convexity of F, i.e. $\frac{d^2}{dx^2}F(x) \leq 0$. It comes up iff:

$$\frac{-\xi + x(1-\xi)}{x^2} \leq 0$$

Therefore

$$x \leq \frac{\xi}{1-\xi}.$$

Since $\frac{\xi}{1-\xi} \geq 1$ for $\xi \geq \frac{1}{2}$ we have $B_j < \tau_j$. In addition if $E(R_j) \geq E(R_0)$, then $\tau_j E(R_j) - B_j E(R_0) > 0$.

Now, it is enough to show that $\forall j$: $\ln A_j(\beta,\xi,1) < 0$ i.e. $A_j(\beta,\xi,1) < 1$, then $\breve{\varrho}^{aux} \geq 0$.

$$A_j(\beta,\xi,1) = \frac{\beta^{(1-\xi)\tau}}{\xi(1-\beta^\tau)+\beta^\tau} < 1. \tag{5.30}$$

Denote $\beta^\tau = x$, then $\forall x \in [0,1]$ the following has to be fulfilled:

$$x^{1-\xi} < \xi(1-x) + x. \tag{5.31}$$

Denote $F(x) = \xi(1-x) + x - x^{1-\xi}$. For $x = 0$ it holds that $F(0) = \xi > 0$, and for $x = 1$, $F(1) = 0$. Since

$$\frac{d}{dx}F(x) = (1-\xi)(1-\frac{1}{x^{\xi}}) < 0$$

we have $F(x) > 0$ for $0 \leq x < 1$. Hence $A_j(\beta, \xi, 1) < 1$.

$\square$

Chapter 6

Internal calibration methods and their results

In the previous chapter we have proposed a new method how the four original parameters of the one factor interest rate models, especially the CIR and the Vašiček model, can be reduced to three new parameters. The idea of reducing the four dimensional parameter space comes on one hand from the paper of Pearson and Sun [56] and on the other hand from the option pricing theory. According to the defined loss functional U which is the difference between the computed and real market yield curve, we are able to find a one dimensional λ-parameterized curve of global minimizers of U. It means that we find three new parameters of the one factor interest rate model depending on the fourth parameter λ the market price of risk.

Our goal is now to propose a method to obtain an exact parameter or an interval of parameters from this one dimensional curve of parameters. All these calculations are implemented on the basis of the assumption that during the calculation only the data are used. This is the reason why this method is called internal.

6.1 Calibration based on maximization of the restricted likelihood function

In this section the calibration method for the estimation of one factor models parameters is discussed. The main principles of the calibration of the CIR and Vašiček model parameters are the same. Let us suppose that the parameter determining the type of the model γ is given.

The method consists of two steps. In the first step, as it was described in Section 5.3, we identify one dimensional curve of the model parameters by minimizing the loss functional. This method looks like the least square approach in linear regression, but the proposed minimization problem is highly non-linear. Having identified the curve of global minimizers of the loss functional we proceed by the second step.

This step consists of maximization of the likelihood function restricted to that curve so the global maximum is attained in a unique point which is the estimation of the model parameters.

Let us propose the second step of our method. Notice that the aim of the first "minimization" step of the method described above was to find a point $(\check{\beta}, \check{\xi}, \check{\varrho})$ - a unique global minimum of the loss functional $U = U(\beta, \xi, \varrho)$. Bearing in mind parameter reduction described in Subsection 5.1 there exists a C^∞ smooth one dimensional curve of original model parameters $(\kappa_\lambda, \theta_\lambda, \sigma_\lambda, \lambda) \in R^4$ parameterized by $\lambda \in \check{J}$ corresponding to the same transformed triple $(\check{\beta}, \check{\xi}, \check{\varrho})$ for which the minimum of U (in terms of transformed variables β, ξ, ϱ) is attained. In order to construct estimation of the model parameters $\kappa, \theta, \sigma, \lambda$ we proceed with the second optimization step in which we find a global maximum of the standard Gaussian likelihood function (LF) over the above mentioned λ-parameterized curve representing of global minimizers of the loss functional U. The two step optimization method combines the maximum likelihood estimation with minimization of the loss functional U. In the case of parameter estimation of a stand-alone short rate process having the form (2.6) the LF is:

$$\ln L(\kappa, \sigma, \theta) = -\frac{1}{2} \sum_{t=2}^{n} \left(\ln v_t^2 + \frac{\varepsilon_t^2}{v_t^2} \right) \tag{6.1}$$

where $v_t^2 = \frac{\sigma^2}{2\kappa} \left(1 - e^{-2\kappa}\right) r_{t-1}^{2\gamma}$, $\varepsilon_t = r_t - e^{-\kappa} r_{t-1} - \theta \left(1 - e^{-\kappa}\right)$ (see [9, 10, 11]). If estimation of model parameters (κ, σ, θ) is realized by maximization of the likelihood function over the whole set $\mathbb{R}_+^3$ then the maximum is unrestricted. The value of the unrestricted maximum likelihood function is:

$$\ln L^u = \ln L(\kappa^u, \sigma^u, \theta^u) = \max_{\kappa, \sigma, \theta > 0} \ln L(\kappa, \sigma, \theta) \, . \tag{6.2}$$

In our approach we make use of restricted maximization of $\ln L$ over the λ-parameterized curve $\{(\kappa_\lambda, \theta_\lambda, \sigma_\lambda), \lambda \in \check{J}\}$. This can be expressed in original model parameters as follows:

$$\ln L^r = \ln L(\kappa_{\bar{\lambda}}, \sigma_{\bar{\lambda}}, \theta_{\bar{\lambda}}) = \max_{\lambda \in \check{J}} \ln L(\kappa_\lambda, \sigma_\lambda, \theta_\lambda), \tag{6.3}$$

where $\check{J} = (-\infty, -(2\check{\xi} - 1) \ln \check{\beta})$ in the case of the CIR model and $\check{J} = \mathbb{R}$ for Vašiček model. The argument $\bar{\kappa} = \kappa_{\bar{\lambda}}$, $\bar{\sigma} = \sigma_{\bar{\lambda}}$, $\bar{\theta} = \theta_{\bar{\lambda}}$ of the maximum of the restricted likelihood function $\ln L^r$ is adopted as a result of two step optimization method for calibrating the model parameters. A global maximizer of the unrestricted likelihood function $\ln L^u$ has been computed by the same variant of the ES algorithm described in Section 5.4. Since maximization of the restricted likelihood function $\ln L^r$ is performed over one dimensional parameter λ and the function $\lambda \mapsto \ln L(\kappa_\lambda, \sigma_\lambda, \theta_\lambda)$ is smooth we could apply a standard optimization software package Mathematica in order to find a global maximizer of the restricted likelihood function. For measuring of accuracy of calibration we introduce the maximum likelihood ratio (MLR) as a ratio of the maximum values of the restricted $\ln L^r$ and unrestricted $\ln L^u$ likelihood functions. We have MLR≤ 1 and if MLR is close to 1 then the restricted maximum likelihood value is close to the unrestricted one. In this case one can therefore expect that the estimated values $(\bar\kappa, \bar\sigma, \bar\theta)$ of the model parameters are close to the argument $(\kappa^u, \sigma^u, \theta^u)$ of the unique global maximum of the unrestricted likelihood function. It may indicate that a simple estimation of parameters based on the mean reversion equation (2.6) for the short rate process r_t is also suitable for estimation of the whole term structure.

6.1.1 Qualitative measure of goodness of fit. Non-linear R^2 ratio

In linear regression statistical methods, the appropriateness of linear regression function is measured by the R^2 ratio. If the value of R^2 ratio is close to one, it indicates that the given data set can be regressed by a linear function. In the case of non-linear regression, there is no unique way how to define the equivalent concept of the linear R^2 ratio. The non-linear R^2 ratio essentially depends on the choice of the reference value. We take this value of the loss functional (5.12) by taking the argument $(\beta, \xi, \varrho) = (1, 1, 1)$. Since $\lim_{\beta\to 1} B_j = \tau_j$ and $\ln A_j = 0$ for $\beta = 1$ it is easy to calculate that

$$U(1,1,1) = \frac{1}{m}\sum_{j=1}^{m} \tau_j^2 E((R_j - R_0)^2),$$

and, moreover, $U(1,1,1) = U(1, \xi, \varrho)$ for any $\xi \in [0,1]$ and $\varrho \in \mathbb{R}$.

Now we are able to define the non-linear R^2 ratio measuring the quality of non-linear regression as follows:

$$R^2 = 1 - \frac{U(\check\beta, \check\xi, \check\rho)}{U(1,1,1)}, \tag{6.4}$$

where $(\check\beta, \check\xi, \check\rho)$ is the argument of the unique global minimum of the loss functional U. Then $0 \le R^2 \le 1$. The value of R^2 close to one indicates perfect matching of the yield curve computed for parameters $(\check\beta, \check\xi, \check\rho)$ and that of the given real market data set.

6.2 Binding of term structures by expected long-term interest rate interval

In the calibration based on the two step optimization method: minimization of the loss functional in the first step and maximization of the restricted likelihood function in the second step we utilize only aggregated statistics from the data. As we can see in the loss functional:

$$\begin{aligned} U(\beta, \xi, \varrho) &= \frac{1}{m}\sum_{j=1}^{m}((\tau_j E(R_j) - B_j E(R_0) + \ln A_j)^2 \\ &+ D(\tau_j R_j - B_j R_0)), \end{aligned}$$

where $E(X_j)$ is the mean and $D(X_j)$ is the dispersion of $X_j = \{X_j^i, i = 1, ..., n\}$. It means that we calculate the mean and the covariance, so we use aggregated time series information from the term structure of interest rates and in the calibration we use only this information. The yield curve comprises more information which can be also utilized during the calibration. So in the second step of the calibration we may also utilize an extra information from the data. It is important that we use only the data basis to get the final parameters. In this method we use the yield curves and try to bound their means or the whole term structure. Using this bounding we get not only a unique point but an interval of parameters.

The main idea of this new approach is the follows:

- in the first step we minimize the loss functional $U = U(\beta, \xi, \varrho)$ on the λ-parameterized curve of global minimizers of U,
- in the second step we use an extra information from the data. We utilize the variety richness of the yield curves to calibrate the fourth parameter λ.

The expected long-term interest rate interval $[\theta_d, \theta_u]$ which is obtained from the data basis, is used in the calibration. In this case we do not get a specific calibrated point of parameters in the four dimensional parameter space of unknown values in the case of the CIR and Vašiček model. Details of the second step are explained in the next two subsections.

6.2.1 Mean value binding

Recall that according to the CIR model the price of a zero-coupon bond $P = P(t, T, r)$ satisfies the parabolic equation (5.1) and is given by the explicit formula $P(t, T, r) = A(T-t)e^{-B(T-t)r}$. Thus $\partial_r P = -BP$. Hence equation (5.1) can be rewritten as

$$\frac{\partial P}{\partial t} + \kappa(\theta - r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r \frac{\partial^2 P}{\partial r^2} - r^* P = 0\,, \quad t \in (0, T)\,,\ r > 0 \tag{6.5}$$

where $r^* = (1 - \lambda B)r$. According to [56] the multiplier $1 - \lambda B$ can be interpreted as the risk premium factor and r^* as the expected rate of return on the bond. Since $B(\tau) > 0$ for $\tau > 0$ we have $r^* > r$ if $\lambda < 0$. On the other hand, if $\lambda > 0$, market bond return r^* is less than risk less return rate r.

In the case of Vašiček model the equation for the risk premium factor is slightly different, because of the form of the function $\tilde{\lambda}(r)$ in (2.12). We have $r^* = r - \lambda\,\sigma B$ for this model.

In the first step of minimization described in Chapter 5 we obtain a point $(\check{\beta}, \check{\xi}, \check{\varrho})$ which is, in the original model, represented by a curve of parameters $(\kappa_\lambda, \theta_\lambda, \sigma_\lambda, \lambda) \in \mathbb{R}^4$ parameterized by $\lambda \in \check{J}$. It means that we need to find λ and then the $(\kappa_\lambda, \theta_\lambda, \sigma_\lambda, \lambda)$ is unambiguously defined. The same curve admits many other parameterization, in particular, parameterization by θ in the form $(\kappa_\theta, \theta, \sigma_\theta, \lambda_\theta) \in \mathbb{R}^4$.

In this part we propose a new way of calibration of the CIR and Vašiček model. This method is based on a targeting interval of the expected long-term interest rate $I_\theta = [\theta_d, \theta_u]$. We can determine an interval for the market price of risk $I_\lambda = [\lambda_d, \lambda_u]$, where $\lambda_d = \lambda(\theta_d)$ and $\lambda_u = \lambda(\theta_u)$, if this interval is known. In this approach we relate:

$$r^* = r^*(\lambda, \tau_j, R_0^i) = \begin{cases} (1 - \lambda B(\tau_j))R_0^i & (CIR) \\ R_0^i - \lambda\sigma B(\tau_j) & (Vasicek) \end{cases} \tag{6.6}$$

where the first part stands for CIR model, and second one stands for Vašiček model. Notice that σ depends only on $(\check{\beta}, \check{\xi}, \check{\varrho})$ for both CIR and Vašiček model. Based on this relationship, for the targeting interval of the expected long-term interest rate $I_\theta = [\theta_d, \theta_u]$ we can define an interval for the expected rate of return as follows:

$$\begin{aligned} r_u^* &= r^*(\lambda(\theta_d), \tau_j, E(R_0)), \\ r_d^* &= r^*(\lambda(\theta_u), \tau_j, E(R_0)), \end{aligned}$$

for both of the models.

The idea of the calibration based on binding of means of yield curves, consists in finding the narrowest possible interval $I_\theta = [\theta_d, \theta_u]$ such that:

$$r_d^* < E(R_j) < r_u^* \quad \forall j \in \{1, ..., m\}.$$

To be more precise, we need to determine the relation between r^* and θ for CIR and Vašiček models.

- Case of the CIR model:
  Based on the parameter reduction presented in Section 5.1 we have derived in (5.6) the λ-parameterized curve in the CIR model case. According to these parameters the mapping $\theta \longmapsto \lambda(\theta)$ is unambiguously defined as follows:

  $$\lambda(\theta) = -\check{K} - \check{\varrho}\,\frac{\sigma^2}{2\theta},$$

  where $\check{K} = (2\check{\xi} - 1)\ln\check{\beta}$ and $\sigma = -\ln\check{\beta}\sqrt{2\check{\xi}(1 - \check{\xi})}$.

- Case of the Vašiček model:
  The λ-parameterized curve for this model is defined in (5.10). In this case the above mentioned mapping is as follows:

$$\lambda(\theta) = \left(\theta - \check{\xi} - \frac{\sigma^2}{2(-\ln\check{\beta})^2}\right)\left(-\frac{\ln\check{\beta}}{\sigma}\right)$$

  where $\sigma = 2\sqrt{(-\ln\check{\beta})\check{\varrho}}$.

Since $\check{\beta} \in (0,1)$, $\check{\varrho} > 0$ and $\sigma^2 > 0$ we have the following:

Proposition 6.2.1. *Function $\theta \longmapsto \lambda(\theta)$ is increasing and $\theta \longmapsto r^*(\lambda(\theta), \tau_j, E(R_0))$ is decreasing in θ for the CIR and Vašiček models. Moreover, $\lambda(0^+) = \infty$, $\lambda(\infty) = -\check{\kappa}$, $r^*(\lambda(0^+), \tau_j, E(R_0)) = \infty$, for CIR model; $\lambda(0^+) < 0$, $\lambda(\infty) = \infty$, $r^*(\lambda(\infty), \tau_j, E(R_0)) = -\infty$, for Vašiček model.*

An illustration can be seen on Figure 6.1. To find the appropriate targeting interval for the expected long-term interest rate we have to solve the following problem, to detect whether the following sets are not empty:

$$\begin{aligned}\theta_u &= \inf\{\theta > 0 \mid \forall j \in \{1...m\}, \quad r^*(\lambda(\theta), \tau_j, E(R_0)) < E(R_j)\} \\ \theta_d &= \sup\{\theta > 0 \mid \forall j \in \{1...m\}, \quad r^*(\lambda(\theta), \tau_j, E(R_0)) > E(R_j)\}\,.\end{aligned} \tag{6.7}$$

We remind ourselves (see (6.6)) that $r^*(\lambda(\theta), \tau_j, E(R_0)) = (1 - \lambda(\theta)B_j)E(R_0)$ in the case of the CIR model, and, $r^*(\lambda(\theta), \tau_j, E(R_0)) = E(R_0) - \lambda(\theta)\sigma B_j$ in the case of the Vašiček model.

Let us consider the CIR model and its mean value binding interval $I_\theta = [\theta_d, \theta_u]$. It follows from Proposition 6.2.1 that $r^*(\lambda(0^+), \tau_j, E(R_0)) = \infty$. If $(1+\check{K}B(\tau_j))E(R_0) \leq \min_{1\leq k\leq m} E(R_k)$, $\forall j \in \{1, ..., m\}$ then $r^*(\lambda(+\infty), \tau_j, E(R_0))) \leq \min_{1\leq k\leq m} E(R_j)$, $\forall j \in \{1, ..., m\}$ and therefore there exist finite values of $0 < \theta_d \leq \theta_u < \infty$ defined as in (6.7). Thus we have shown the following proposition:

Proposition 6.2.2. *Suppose that $(1+\check{K}B(\tau_j))E(R_0) \leq \min_{1\leq k\leq m} E(R_k)$, $\forall j \in \{1, ..., m\}$. Then the mean value binding interval $[\theta_d, \theta_u]$ for the CIR model is finite, i.e. $0 < \theta_d \leq \theta_u < \infty$. This condition is fulfilled if $\check{K} < 0$ i.e. $\check{\xi} > \frac{1}{2}$ and $E(R_0) \leq E(R_j)$ $\forall j \in \{1, ..., m\}$.*

6.2.2 Binding of the whole term structures and their significant parts

In the previous part the mean value have been bound. We are also interested in whether it is possible to bind (of course over a reasonable interval) not only the mean value of the yield curves by $[r_d^*, r_u^*]$ but also the whole term structure or their

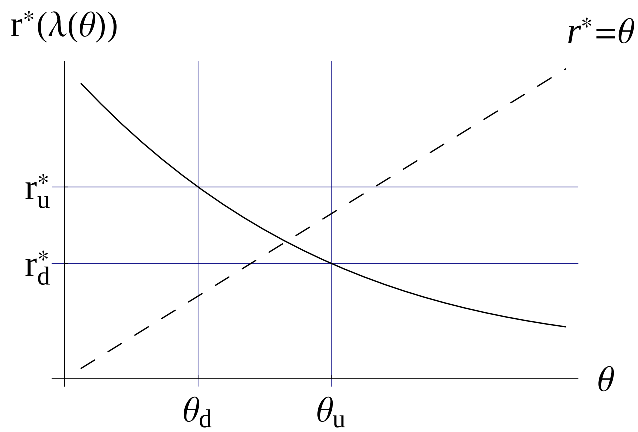


Figure 6.1: Illustration: where $r^*(\lambda(\theta)) = r^*(\lambda(\theta), \tau_j, E(R_0))$

significant parts (e.g. 95%) through the targeting interval of expected long-term interest rate $I_\theta = [\theta_d, \theta_u]$. The new task is to find this interval as the following sets are not empty:

$$\begin{aligned} \theta_u &= \inf\{\theta > 0,\ Pr(j \in \{1...m\} \mid r^*(\lambda(\theta), \tau_j, R_0^i) < R_j^i,\ \forall i \in \{1...n\}) \geq 1 - \alpha\} \\ \theta_d &= \sup\{\theta > 0,\ Pr(j \in \{1...m\} \mid r^*(\lambda(\theta), \tau_j, R_0^i) > R_j^i,\ \forall i \in \{1...n\}) \geq 1 - \alpha\} \end{aligned}$$

With this choice of $[\theta_d, \theta_u]$ interval we have the estimate:

$$Pr(j \in \{1...m\} \mid \quad r^*(\lambda(\theta_u), \tau_j, R_0^i) < R_j^i < r^*(\lambda(\theta_d), \tau_j, R_0^i),\ i \in \{1...n\}) \geq 1 - 2\alpha,$$

where Pr is the probability of the event and α is the significance or confidence level.

If $1 - \alpha = 1$, i.e. $\alpha = 0$, then, for all $i = 1, .., n$ and $j = 1, ..., m$, we have:

$$r^*(\lambda(\theta), \tau_j, R_0^i) < R_j^i$$

and therefore we have

$$r^*(\lambda(\theta), \tau_j, E(R_0)) < E(R_j)$$

because r^* is linear in R_0^i argument (see 6.6). It means that $\theta_u^{\alpha=0} > \theta_u^{mean}$ [1] obtained from mean value binding. On the other hand, the same can be derived for the θ_d, so we have that $\theta_d^{\alpha=0} < \theta_d^{mean}$.

6.3 Internal calibration results for a sample European term structures

The aim of this section is to present results of the two step optimization method for terms structures for various European countries and to make comparison of stable

[1] $\theta_u^{\alpha=0}$ stands for θ_u when $\alpha = 0$ and θ_u^{mean} stands for θ_u when the parameter is calculated based on mean value binding.

western European financial markets to those of the new EU member states. More precisely, we will compare European inter-bank offered rates EURIBOR, London inter-bank offered rates EURO-LIBOR and USD-LIBOR, and inter-bank offered rates in Slovakia (BRIBOR), Hungary (BUBOR), the Czech Republic (PRIBOR) and Poland (WIBOR).

6.3.1 Term structure description and summary statistics

In the two step optimization method for calibration of the CIR and Vašiček model we need to have a good approximation for the process of the spot interest rate (2.6) for inter-bank loans. The overnight or short rate seems to be appropriate. The data for the short interest rate are available for all investigated financial markets (BRIBOR, PRIBOR, WIBOR, BUBOR) as well as for London inter-bank offered rates (LIBOR). In case of the Euro-zone term structure EURIBOR, there exists an accepted substitute to the overnight called EONIA. It is computed as a weighted average of all overnight unsecured lending transactions in the inter-bank market, initiated within the Euro area by the Panel Banks. These Panel Banks (the overall number 48) are the same as the banks for the quoting of EURIBOR. They use the convention of 360 day count and the inter-bank rate is displayed to two decimal places. In the case of BRIBOR the number of the Panel Banks is 7, the rates, quoted at 11:00 a.m., are computed as non-weighted arithmetic average of all quoted values except the lowest and the highest one and displayed to two decimal places. The number of Panel Banks for BUBOR is 16, for WIBOR is 10 and the rates are quoted at 10:30 a.m. in the first case and at 11:00 a.m. in the second.

Term structures for the above mentioned European financial markets data contain bonds with the following maturities:

- EURO-LIBOR and USD-LIBOR: 1 week, 1 up to 12 months, i.e. its length is $m = 13$.
- EURIBOR: 1, 2 and 3 weeks, 1 up to 12 months, i.e. its length is $m = 15$.
- BRIBOR: 1 and 2 weeks and 1, 2, 3, 6, 9, 12 months, i.e. its length is $m = 8$.
- BUBOR: 1 and 2 weeks, 1, 3, 6, 9, 12 months, i.e. its length is $m = 7$.
- WIBOR 1 week and 1, 3, 6, 9, 12 months, i.e. its length is $m = 6$.
- PRIBOR: 1 and 2 weeks, 1, 2, 3, 6, 9, 12 months, i.e. its length is $m = 8$.

The sample mean and the standard deviation for different inter-bank offered rates is presented on quarterly basis in Table 6.1 for year 2003. The mean of BRIBOR, WIBOR and BUBOR is higher than the mean of the last three data sample (PRIBOR, EURIBOR and EURO-LIBOR) during the whole year. The same is true concerning the standard deviation. It is an indication that the Czech data could have similar qualitative properties as the western European term structures.

Table 6.1: Descriptive statistics for various term structures. Mean and standard deviation (STD) (in %) are shown for the overnight rate and the rate on the longest bond with 1-year maturity.

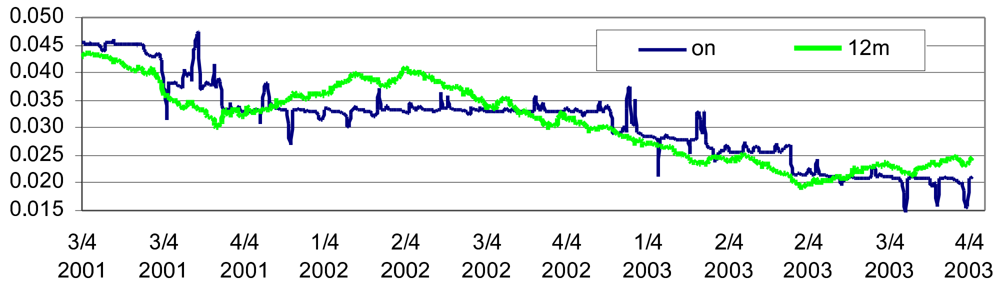
		1/4 2003		2/4 2003		3/4 2003		4/4 2003	
		Mean	STD	Mean	STD	Mean	STD	Mean	STD
BRIBOR	on	5.75	1.041	6.27	1.279	5.63	0.802	5.48	0.992
	1y	5.48	0.205	5.42	0.208	5.80	0.066	5.50	0.033
WIBOR	on	6.65	0.761	5.76	0.359	5.22	0.438	5.17	0.438
	1y	5.95	0.138	5.19	0.255	4.97	0.053	5.79	0.380
BUBOR	on	5.42	1.813	7.08	0.879	9.58	0.547	10.52	1.213
	1y	6.57	0.433	6.76	0.773	8.80	0.207	10.02	1.334
PRIBOR	on	2.52	0.107	2.44	0.045	2.06	0.135	1.94	0.032
	1y	2.43	0.130	2.33	0.084	2.13	0.063	2.19	0.061
EURIBOR	on	2.77	0.188	2.44	0.199	2.07	0.120	2.02	0.169
	1y	2.54	0.140	2.23	0.189	2.20	0.106	2.36	0.081
EURO-LIB	on	2.79	0.196	2.47	0.196	2.08	0.101	2.02	0.139
	1y	2.54	0.139	2.23	0.187	2.20	0.105	2.35	0.085

The key issue of the CIR and Vašiček model is to model the short rate as an Orstein-Uhlenbeck type of the mean reverting process (2.6). In Figure 6.2 we present some examples of short rate for specific samples of data (EURO-LIBOR, BRIBOR and PRIBOR). As we can see the over-night is more volatile than the interest rates with longer maturity.

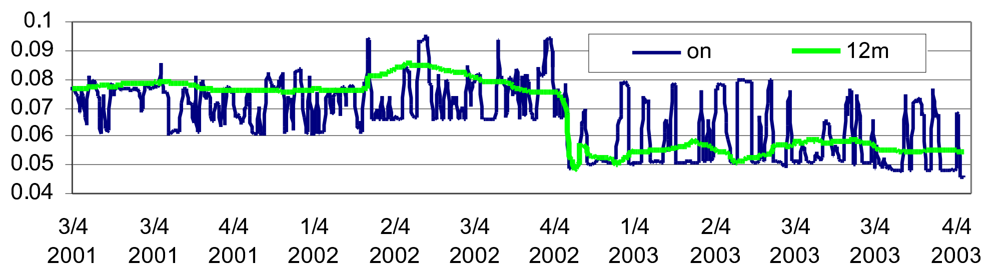
6.3.2 Results of calibration by means of restricted likelihood function

The results of calibration for the CIR model parameters as well as corresponding maximum likelihood (MLR) and R^2 ratios are summarized in Table 6.2 for term structures with maturities up to one year.

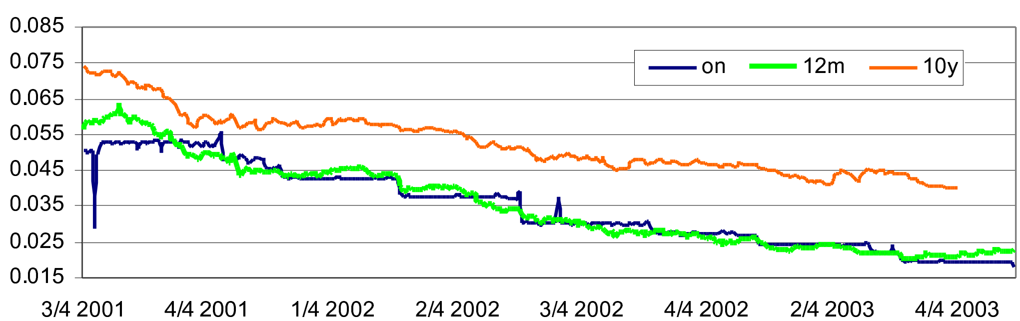
Table 6.2 reports quarterly results for BRIBOR, WIBOR, BUBOR, PRIBOR, EURO-LIBOR and EURIBOR for the year 2003. Estimated parameters $\kappa, \sigma, \theta, \lambda$, the value of the loss functional (U) and the non-linear R^2 ratio together with the maximum likelihood ratio (MLR) are presented. Behavior of the expected long-term interest rate θ is in accordance with the expectancy of the market in the long-term run. It predicts interest rates close to 1.7% for EURO-LIBOR and EURIBOR as well as for PRIBOR. Other term structures also indicate decrease of interest rates in the future but these estimations of θ are quantitatively less convincing compared to EURO-LIBOR, EURIBOR and PRIBOR predictions. Results of estimation for κ show that the speed of reversion for EURIBOR and EURO-LIBOR is comparable. The lowest values of estimated κ were achieved by the PRIBOR term structure and the highest values were achieved for BRIBOR which is in accordance to highly fluctuating character of



(a)



(b)



(c)

Figure 6.2: Graphical description of overnight (short-rate) interest rates and those of bond with longer maturity. Daily data are plotted for EURO-LIBOR (a), BRIBOR (b) and PRIBOR (c). The 10y PRIBOR stands for the 10 year yield on government bonds.

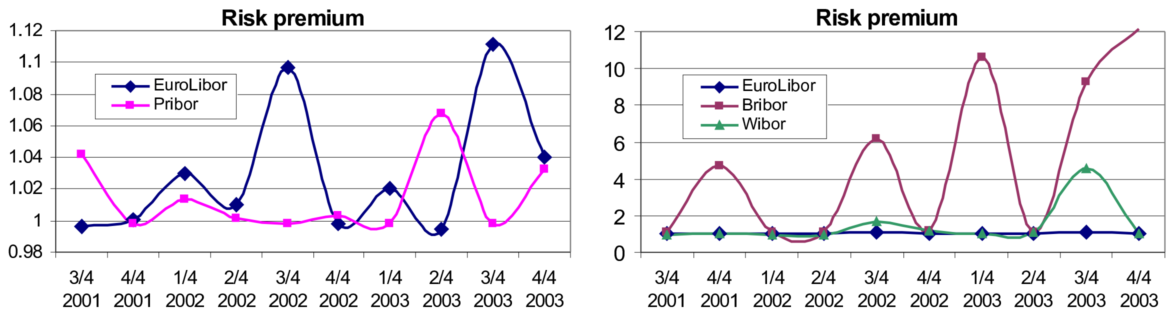


Figure 6.3: Comparison of risk premium factors $1 - \lambda B_1$ for EURO-LIBOR and BRIBOR (left) and EURO-LIBOR, WIBOR and BRIBOR (right).

BRIBOR interest rates (see Figure 6.2). It can be explained by the characteristics of Slovak inter-bank market. The estimated volatility parameter σ of the mean reverting process has the same behavior for WIBOR and BUBOR, and, it is very large for the Slovak BRIBOR data. On the other hand, results for Czech data enables us to conclude that volatility of PRIBOR quantitatively and qualitatively is very similar to that of EURIBOR and EURO-LIBOR. In terms of the maximum likelihood ratio, measuring appropriateness of the CIR model, the overall quality of estimation is better for PRIBOR, EURIBOR and EURO-LIBOR. The non-linear R^2 ratio is mostly close to one for all data, but there are some exceptions.

In Figure 6.4, parts (a) and (b) we show comparison of the MLR and R^2 ratios for EURO-LIBOR an USD-LIBOR. As we can see they are strongly correlated. Hence, in the next parts of this figure, we present a comparison of results of parameter estimation for EURO-LIBOR, BRIBOR, PRIBOR and WIBOR only. The graph (c) displays their MLR. In most quarters, this ratio is better for EURO-LIBOR and the worst for BRIBOR. For the last two samples (PRIBOR and WIBOR), it is varying. In the graph (d) we can see that the best R^2 ratio is achieved for WIBOR and the worst for PRIBOR. This value is varying for BRIBOR and EURO-LIBOR. Parts (e) and (f) present estimated parameters θ (expected long-term interest rate) and σ (volatility of the process (2.6)). For PRIBOR and EURO-LIBOR the results are similar not only for θ but also for σ. The volatility of the process for Slovak data is very high and also parameter θ is quite volatile.

Risk premium analysis

In this section we discuss and analyze results of parameter estimation for the parameter λ representing the market price of risk in the CIR model.

Table 6.2 shows that the market price of risk λ is negative in most time periods. It implies that the expected rate of bond return r^* is greater than instantaneous rate r. There are however some short time periods in which the market price of risk is positive for EURIBOR and EURO-LIBOR (2nd quarter). In Figure 6.3 we plot the

Table 6.2: Numerical results of calibration for short term structures (up to one year) for BRIBOR, WIBOR, BUBOR, PRIBOR, EURIBOR and EURO-LIBOR. Results cover 4 quarters of 2003.

BRIBOR	κ	σ	θ	λ	U ($\times 10^{-6}$)	R^2	ML ratio
1/4 2003	688.298	8.960	0.0025	-658.657	1.629	0.947	0.528
2/4 2003	38.467	1.509	0.0458	-6.302	1.704	0.971	0.719
3/4 2003	598.875	8.276	0.0031	-568.839	0.487	0.971	0.536
4/4 2003	793.487	9.396	0.0022	-764.117	0.339	0.986	0.551
WIBOR							
1/4 2003	10.103	0.622	0.0564	-0.362	0.660	0.966	0.702
2/4 2003	7.459	0.877	0.0204	-4.457	1.075	0.897	0.519
3/4 2003	193.565	6.097	0.0029	-189.986	1.781	0.870	0.371
4/4 2003	2.910	0.842	0.0004	-3.084	22.862	0.048	0.388
BUBOR							
1/4 2003	14.233	0.576	0.0555	-2.425	6.118	0.965	0.964
2/4 2003	14.166	0.671	0.0626	-0.658	2.967	0.805	0.684
3/4 2003	134.256	5.064	0.0035	-130.883	1.037	0.940	0.363
4/4 2003	2.715	0.812	0.0111	-2.508	19.593	0.208	0.394
PRIBOR							
1/4 2003	0.098	0.007	0.0248	0.092	0.134	0.633	0.904
2/4 2003	36.934	0.728	0.0209	-4.714	0.238	0.428	0.514
3/4 2003	2.823	0.060	0.0201	-0.200	0.028	0.897	0.814
4/4 2003	3.385	0.097	0.0187	-0.626	0.088	0.924	0.685
EURIBOR							
1/4 2003	40.927	1.030	0.0202	-8.724	0.506	0.783	0.654
2/4 2003	0.818	0.028	0.0240	0.252	0.319	0.746	0.876
3/4 2003	39.209	0.644	0.0178	-6.986	0.143	0.807	0.735
4/4 2003	15.592	0.360	0.0180	-3.451	0.145	0.941	0.778
EURO-LIBOR							
1/4 2003	34.118	0.689	0.0241	-1.897	0.634	0.818	0.712
2/4 2003	0.734	0.024	0.0244	0.276	0.432	0.790	0.864
3/4 2003	40.018	0.699	0.0175	-7.797	0.221	0.714	0.703
4/4 2003	9.217	0.286	0.0178	-2.243	0.221	0.929	0.758

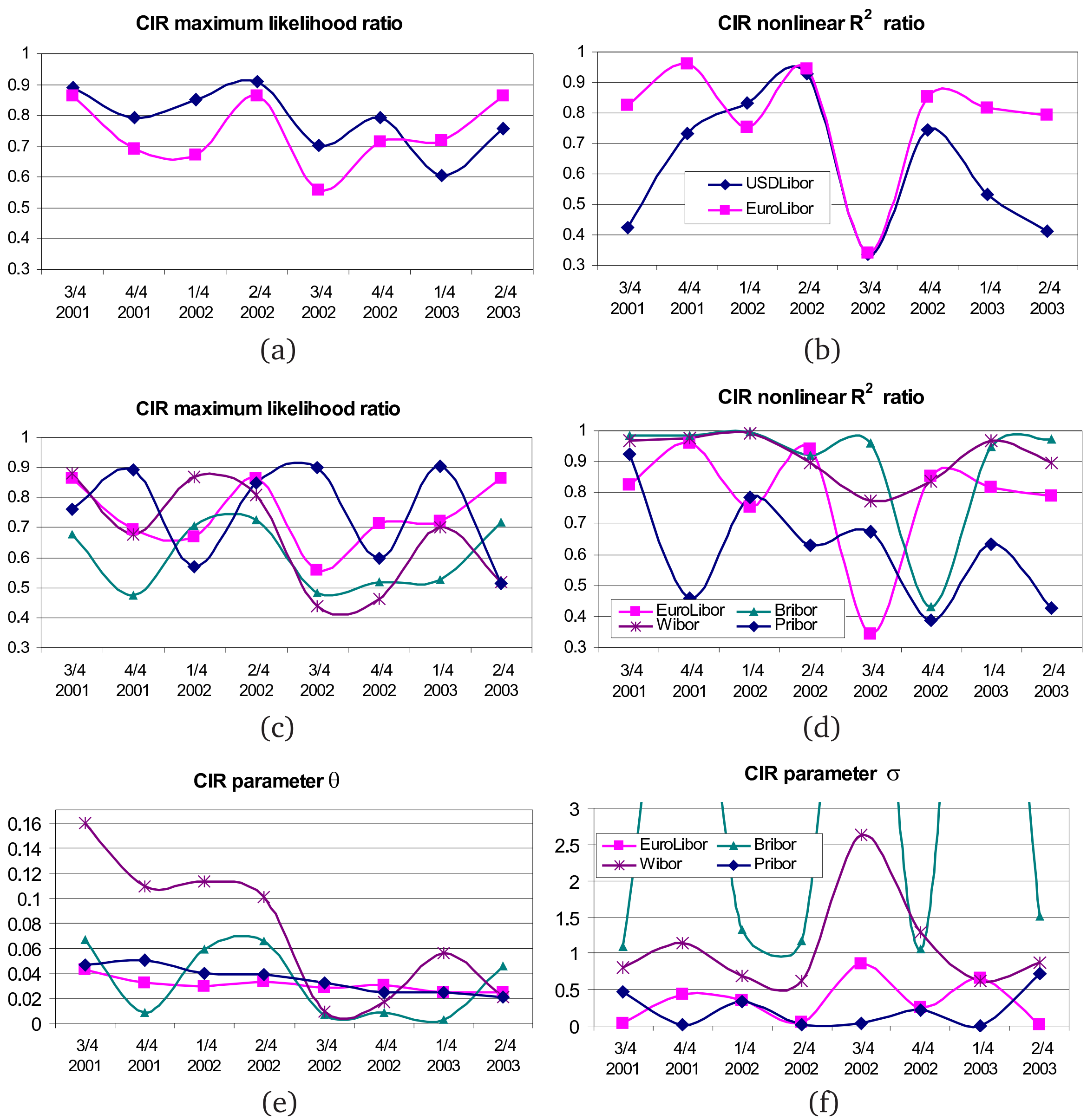


Figure 6.4: Results of parameter estimation for various term structures. Maximum likelihood ratio (a) and R^2 ratio (b) for EURO-LIBOR and USD-LIBOR. Comparison of the same factors for EURO-LIBOR, BRIBOR, PRIBOR and WIBOR is presented in (c) and (d), resp. Estimated parameters θ and σ are shown in (e) and (f), resp.

risk premium factor $1 - \lambda B$ for 10 quarters since the third quarter of 2001. We chose $B = B_1$, i.e. we plotted the risk premium for bonds with one week maturity. Figure 6.3 (left) displays the risk premium of EURO-LIBOR and PRIBOR. They are comparable as far as behavior and range of values are concerned. The right figure presents a comparison for EURO-LIBOR with WIBOR and BRIBOR. The risk premium is quite similar for WIBOR and EURO-LIBOR except of the third quarter of 2003. However, this factor is extremely large and highly volatile for the Slovak BRIBOR term structure.

6.3.3 Results of calibration based on binding of means

Recall that in the mean value binding approach we are seeking for a narrowest interval of the expected long-term interest rate $I_\theta = [\theta_d, \theta_u]$ through the yield curve binding as follows:

$$r_d^* < E(R_j) < r_u^* \quad \forall j \in \{1, .., m\}. \tag{6.8}$$

It is important to remind that this investigation is possible by means of the market price of risk which is different for the Vašiček and CIR models. This is because of the different definition of the expected rate of return of the bond for these models. Based on Sections 5.1.1, 5.1.2, 6.1.1 we have the following formula for the market price of risk depending on the expected long-term interest rate interval:

- case of the Vašiček model:

$$\lambda_i = \left(\theta_i - \check{\xi} + \frac{2\check{\varrho}}{\ln \check{\beta}}\right) \frac{\sqrt{-\ln \check{\beta}}}{2\sqrt{\check{\varrho}}} \tag{6.9}$$

- case of the CIR model:

$$\lambda_i = -\ln \check{\beta} \left[(2\check{\xi} - 1) + \frac{\check{\varrho}}{\theta_i}\check{\xi}(1 - \check{\xi}) \ln \check{\beta}\right] \tag{6.10}$$

where $i \in \{d, u\}$.

According to this fact the algorithm consists of the following steps:

1. obtain $(\check{\beta}, \check{\xi}, \check{\varrho})$ in the first step (minimization of the loss functional)
2. get the narrowest possible interval $I_\theta = [\theta_d, \theta_u]$ in the second step
   - I. give initial approximation $\theta_d \in [\theta_{min}, \theta_{max}]$ and $\theta_u \in [\theta_{min}, \theta_{max}]$ and small steps $\Delta\theta_u, \Delta\theta_d > 0$
   - II. calculate λ_d, λ_u and consequently r_d^* and r_u^*
   - III. verify $r_u^* \leq E(R_j) \leq r_d^*$ for all $j \in \{1, ..., m\}$ (in a special case just for a given subset of $j \in \{1, ..., m\}$)
   - IV. if III is satisfied increase $\theta_d = \theta_d + \Delta\theta_d$ and decrease $\theta_u = \theta_u - \Delta\theta_u$. Go to step II otherwise stop.

It means that we identify the smallest interval of the expected long-term interest rate in iterative process. If all the maturities have to fulfill the condition (6.8) then the interval is wider than in the special case (when only the longer maturities have to fulfill this condition). Tables 6.3-6.6 summarize the results for PRIBOR and BRIBOR interest rates from February up to September 2005 for the CIR model. This computation is rather general, so it could be performed also for other data basis and for the Vašiček model. These results include not only the interval $I_\theta = [\theta_d, \theta_u]$ but also

Table 6.3: Results for PRIBOR interest rates for data from 2005 in the case of including all maturities

PRIBOR	θ_d	θ_u	λ_d	λ_u
February	0.01005	0.02285	-0.46189	0.00570
March	0.02035	0.02215	-0.24978	0.75840
April	0.00865	0.02390	-0.71028	-0.03909
May	0.01580	0.01810	-0.35461	-0.06380
Jun	0.01705	0.01760	-0.13371	-0.00915
July	0.01660	0.01885	-0.17155	-0.06187
August	0.01700	0.01870	-0.25981	-0.11381
September	0.01560	0.02070	-0.24073	-0.08629

Table 6.4: Results for PRIBOR interest rates for data from 2005 in the case of including only longer maturities (from 2 months upto 1 year).

PRIBOR	θ_d	θ_u	λ_d	λ_u
February	0.02090	0.02285	-0.02713	0.00570
March	0.02190	0.02215	0.67572	0.75840
April	0.02010	0.02390	-0.10987	-0.03909
May	0.01795	0.01810	-0.06961	-0.06379
Jun	0.01750	0.01750	-0.00915	-0.00915
July	0.01860	0.01885	-0.06870	-0.06187
August	0.01840	0.01870	-0.13068	-0.11381
September	0.01970	0.02070	-0.10832	-0.08629

the interval for the market price of risk $I_\lambda = [\lambda_d, \lambda_u]$. While in the case of including only the longer maturities into the computation we can obtain a narrower interval, in some cases even a unique point (see Table 6.4 Jun), in the case of covering the whole term structure we obtain a wider interval of parameters (θ and λ).

Some representative examples are shown on Figures 6.5-6.7. Naturally the yield curve bounds (r_d^* and r_u^*) and the mean value of term structure R_j could be compared on these figures. The results are organized as follows: in the first line there are results for the BRIBOR and in the second line there are for PRIBOR interest rates (for a specified time period, in this case May, July and August); the first column is for all maturities and the second is only for the longer maturities, especially for 2 months up to 1 year.

Calibration based on the data basis (it means that the expected long-term interest rate interval I_θ is obtained from the data) results in a positive outcome from the point of view of negative market price of risk interval I_λ. As we can see in Tables 6.3 and 6.5 in the most of the cases $\lambda_u < 0$. The same result is achieved in the case of

Table 6.5: Results for BRIBOR interest rates for data from 2005 in the case of including all maturities

BRIBOR	θ_d	θ_u	λ_d	λ_u
February	0.02805	0.04090	-1.29012	-0.34379
March	0.00580	0.03210	-7.04533	-0.26786
April	0.01170	0.02905	-10.33360	-0.67071
May	0.02545	0.02880	-0.46294	1.40862
Jun	0.02225	0.02795	-3.64359	0.31737
July	0.01955	0.03030	-8.23568	-1.33360
August	0.02825	0.02995	-0.16431	0.70541
September	0.02270	0.02745	-4.26512	-0.81113

Table 6.6: Results for BRIBOR interest rates for data from 2005 in the case of including only longer maturities (from 2 months upto 1 year).

BRIBOR	θ_d	θ_u	λ_d	λ_u
February	0.03285	0.04090	-0.84777	-0.34379
March	0.02400	0.03210	-0.77124	-0.26786
April	0.02535	0.02905	-1.60358	-0.67071
May	0.02700	0.02880	0.48575	1.40862
Jun	0.02595	0.02795	-0.83528	0.31737
July	0.02805	0.03030	-2.30550	-1.33360
August	0.02860	0.02995	0.03424	0.70541
September	0.02600	0.02745	-1.68687	-0.81113

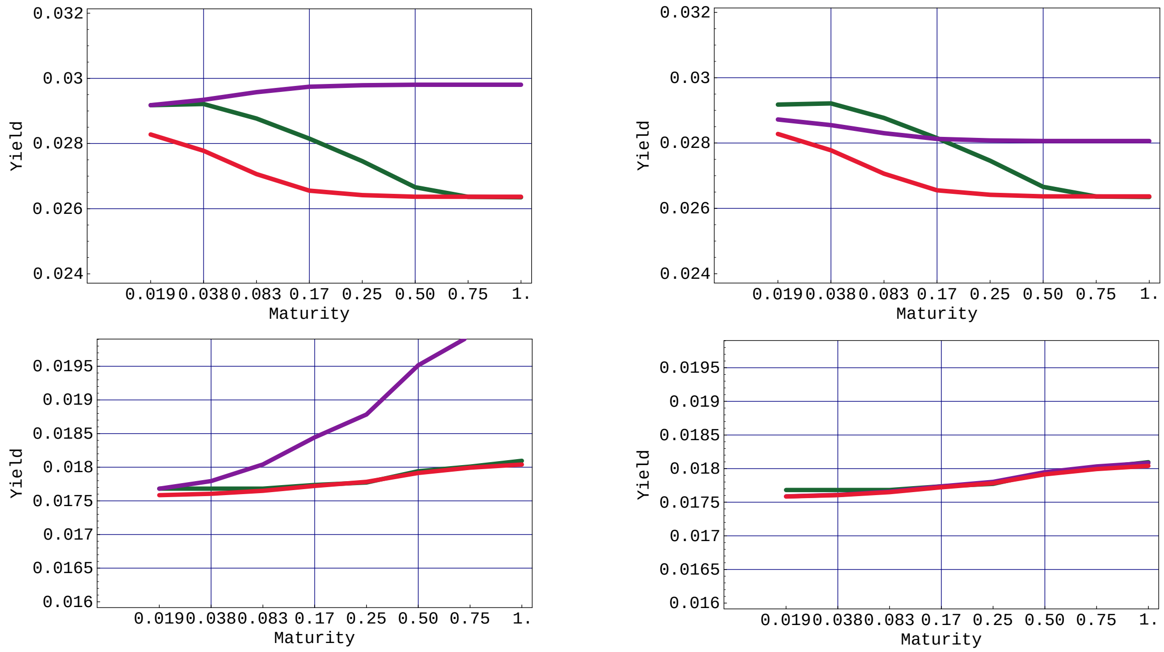


Figure 6.5: Results from May 2005 for BRIBOR (top row) and PRIBOR (bottom row) interest rates (left–all maturities, right–2m up to 1y); r_u^*–violet, r_d^*–red, $E(R_j)$–green

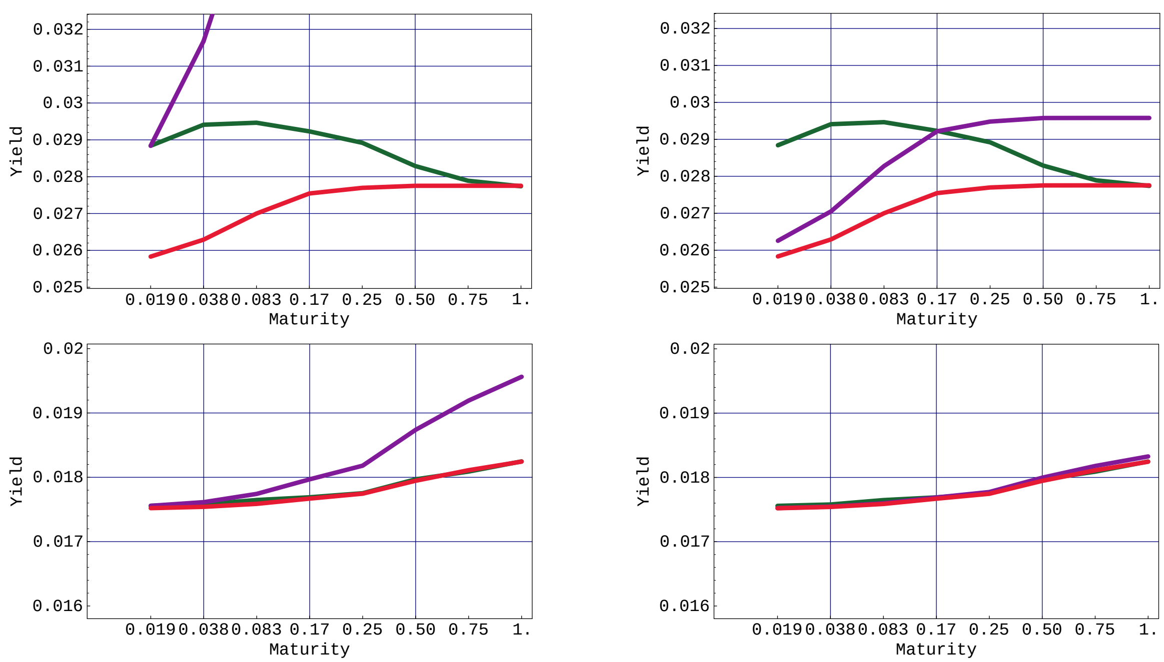


Figure 6.6: Results from July 2005 for BRIBOR (top row) and PRIBOR (bottom row) interest rates (left–all maturities, right–2m up to 1y); r_u^*–violet, r_d^*–red, $E(R_j)$–green

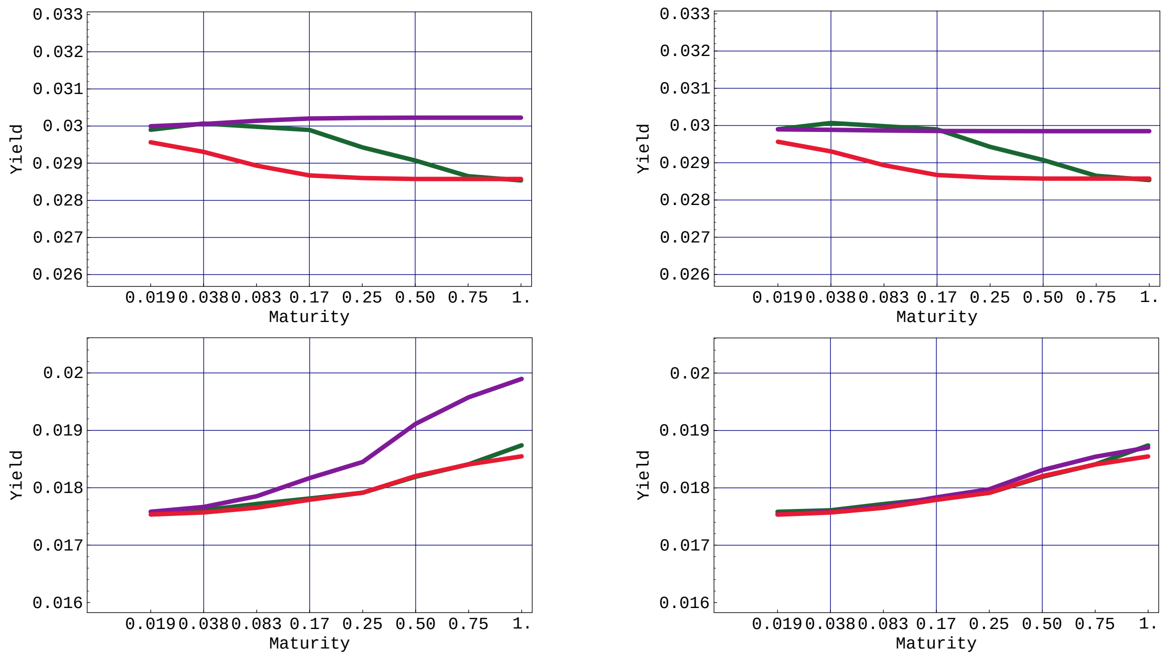


Figure 6.7: Results from August 2005 for BRIBOR (top row) and PRIBOR (bottom row) interest rates (left–all maturities, right–2m up to 1y); r_u^*–violet, r_d^*–red, $E(R_j)$–green

including only the longer maturities into the calibration.

In the case of binding of the whole term structure this approach will be modified in the following way:

1. obtain $(\check{\beta}, \check{\xi}, \check{\varrho})$ in the first step (minimization of the loss functional)

2. get the narrowest possible interval $[\theta_d, \theta_u]$ in the second step

   I. give initial approximation $\theta_d \in [\theta_{min}, \theta_{max}]$ and $\theta_u \in [\theta_{min}, \theta_{max}]$ and small steps $\Delta\theta_d, \Delta\theta_u > 0$

   II. calculate λ_d, λ_u and consequently $r^*(\lambda(\theta_u), \tau_j, R_0^i)$ and $r^*(\lambda(\theta_d), \tau_j, R_0^i)$

   III. verify $r^*(\lambda(\theta_u), \tau_j, R_0^i) \leq R_j^i \leq r^*(\lambda(\theta_d), \tau_j, R_0^i)$ for all $j \in \{1, ..., m\}$ and for all $i \in \{1, ..., n\}$

   IV. if III is satisfied increase $\theta_d = \theta_d + \Delta\theta_d$ and decrease $\theta_u = \theta_u - \Delta\theta_u$. Go to step II otherwise stop.

Chapter 7

External calibration methods

In this chapter we introduce a new approach in the one factor interest rate models calibration. Until now we have tried to calibrate all the unknown parameters of the Vašiček and CIR models on the basis of the inter-bank rates. We searched for the best solution to fit the real data. This method leads to an endogenous outcome. The new approach, proposed in this part, uses not only the inter-bank rates but also an expert expectation. It results in an exogenous solution.

We recall main steps of the option pricing Black-Scholes model in order to understand our motivation for introduction of the external calibration method based on an extra (externally provided) information. We should find out the analogy of the option pricing with the bond pricing theory.

The evolution of the asset price S at time t is assumed to follow the process of the form:

$$dS = \mu S dt + \sigma S dw, \tag{7.1}$$

where μ is the expected rate of return, σ is the volatility and w is the standard Wiener process. According to the assumptions of the risk-less hedging principle (see [47]) one can derive the Black-Scholes partial differential equation for option pricing in the form:

$$\frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0. \tag{7.2}$$

Note that the parameter μ which is the expected rate of return does not enter the above PDE. It indicates that the risk preference of the investor (tendency of stock evolution) do not affect the option price so it exactly depends on three parameters σ, r and E: $V(S, t = T - \tau; \sigma, r, E)$, where E is the strike price of the underlying asset.

The knowledge of the option price V is not enough for the calibration of μ, so it must be externally added.

In the case of the CIR model (the same analogy could be used also for the Vašiček model), the underlying equation is of the form:

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dw_t \tag{7.3}$$

where $\kappa(\theta - r_t)$ is the mean reverting drift and $\sigma\sqrt{r_t}$ is the volatility of the process for r_t. The method of applying the risk-less hedging principle is similar but slightly different from that used in the option pricing model (see [47]). Since the interest rate is not a traded asset, dependence of the zero-coupon bond price on the investors preferences could not be eliminated as it has been done in the case of options. We obtain the following PDE for the price of a zero-coupon bond:

$$\frac{\partial P}{\partial t} + (\kappa(\theta - r) - \lambda r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 r\frac{\partial^2 P}{\partial r^2} - rP = 0. \tag{7.4}$$

In this case the bond price depends on four parameters $P(r, t = T - \tau; \kappa, \sigma, \theta, \lambda)$, but we are able to reduce the four dimensional parameter space of $(\kappa, \sigma, \theta, \lambda)$ into three dimensional space as it was mentioned in Section 5.1. So the bond price depends exactly on three parameters and the fourth one should be determined either internally (using e.g. restricted maximum likelihood estimation) or it can be provided as an exogenous parameter. As we can see there is an analogy with option pricing, the only difference is in the investors preferences, the tendency of evolution of the underlying equation. In the case of the option pricing, this parameter have to be provided, because it could not be identified from the data. The bond price knowledge allows us the calibration of this parameter internally, as it was done in Chapter 6, but the motivation from the option pricing theory suggests us a new approach.

This methodology, the so-called external calibration method, is based on the idea of externally provided parameter. This parameter could be for example the expected long-term interest rate θ. This input can be obtained by an expert data analysis or from the market experts expectations.

Generally, in the one factor interest rate models, there are four parameters involved. According to the economic interpretation of these parameters: the expected long-term interest rate, the market price of risk, the volatility of the short rate process and the mean reversion factor; it is difficult to obtain externally one of them. The most likely to get an expectation about the long-term interest rate parameter θ.

7.1 Calibration based on prescribed expected long-term interest rate interval

Suppose that we are given an extra information that the expected long-term interest rate θ lies within the interval $I_\theta = [\theta_d, \theta_u]$, where θ_d is the lower and θ_u is the upper bound provided by external expertise and $0 < \theta_d < \theta_u$.

Table 7.1: Expected long term interest rate intervals for EURIBOR in 2006.

	θ_d		θ_u	
	on (%)	1w (%)	on (%)	1w (%)
February	2.3299	2.3588	2.3691	2.3862
March	2.3898	2.5656	2.6502	2.6402
April	2.5967	2.6212	2.6600	2.6478
May	2.5162	2.6105	2.6375	2.6318
Jun	2.5365	2.6785	2.8589	2.9067
July	2.7923	2.8310	2.8353	2.8551
August	2.8269	2.9833	3.1096	3.1208
September	3.0159	3.0659	3.0660	3.0820

This problem could be very complicated for multi factor models, because we have to hit the target interval with more parameters than in the case of one factor interest rate models. For simplicity, let us assume a one factor interest rate model for the zero-coupon bond price. As we know, we have an explicit solution to the PDE for CIR and Vašiček model, too. It means that we are able to calibrate the parameters of the models, of course with additional information which is the expected long-term interest rate interval I_θ.

The basic step before the calibration is again the parameter reduction as described in Section 5.1. We can minimize the loss functional (5.13) to obtain the optimum vector $(\check{\beta}, \check{\xi}, \check{\varrho})$. In the second step we need extra information which is the prescribed interval I_θ in our methodology. Based on this expectation we are able to find:

1. the market price of risk interval $I_\lambda = [\lambda_d, \lambda_u]$, where $\lambda_d < \lambda_u$ are defined in (6.9) for the Vašiček model and in (6.10) for the CIR model, i.e.

   CIR model:
$$\lambda_d = -\ln\check{\beta}\left[(2\check{\xi}-1) + \frac{\check{\varrho}}{\theta_d}\check{\xi}(1-\check{\xi})\ln\check{\beta}\right]$$
$$\lambda_u = -\ln\check{\beta}\left[(2\check{\xi}-1) + \frac{\check{\varrho}}{\theta_u}\check{\xi}(1-\check{\xi})\ln\check{\beta}\right]$$

   Vašiček model:
$$\lambda_d = \left(\theta_d - \check{\xi} + \frac{2\check{\varrho}}{\ln\check{\beta}}\right)\frac{\sqrt{-\ln\check{\beta}}}{2\sqrt{\check{\varrho}}}$$
$$\lambda_u = \left(\theta_d - \check{\xi} + \frac{2\check{\varrho}}{\ln\check{\beta}}\right)\frac{\sqrt{-\ln\check{\beta}}}{2\sqrt{\check{\varrho}}}$$

   Notice that $0 < \beta < 1 \Longrightarrow -\ln\beta > 0$.

2. the interval for the speed of reversion $I_\kappa = [\kappa_d, \kappa_u]$ in the CIR model case, where κ_d and κ_u are defined as follows:

$$\kappa_d = -(2\check{\xi} - 1)\ln\check{\beta} - \lambda_u$$

$$\kappa_u = -(2\check{\xi} - 1)\ln\check{\beta} - \lambda_d$$

For the Vašiček model this parameter depends only on the transformed parameters (5.7). $\kappa = \kappa_d = \kappa_u = -\ln\check{\beta}$.

3. and the last parameter of the one factor model σ which is independent on λ in both cases and can be expressed as:

$$\begin{aligned} \sigma = \sigma_d = \sigma_u &= -\ln\check{\beta}\sqrt{2\check{\xi}(1-\check{\xi})} \quad (CIR) \\ \sigma = \sigma_d = \sigma_u &= 2\sqrt{(-\ln\check{\beta})\check{\varrho}} \quad (Vasicek) \end{aligned}$$

7.1.1 Discussion on the expected long-term interest rate interval with results

As it was mentioned earlier, considering an observable or externally provided parameter for the interest rate models, the most rational is to make some expectations about the long-term interest rate θ. However setting up this parameter is difficult. It would be careless to say that this expectation is for sure a fixed value. For that reason, we decided to set an interval for the expected long-term interest rate.

We realize that our targeting need not be right, however the background of this consideration could be plausible. The Slovak economy is a small open economy. We are influenced by other European countries. Our currency is linked to the Euro (in a standard fluctuation band), as we have entered the ERM II in November 2005. We plan to adopt the Euro and enter the Euro-zone in the close future. We have to fulfill the Maastrichts criteria before entering. One of them is about the long term interest rate on 10 year government bonds, that it could not exceed the average interest rate of three EU countries (with the smallest inflation) more than 2%. The central bank will foster the growth of economy through lower interest rates. All these facts influence the market expectations. To avoid the shock from these changes the market is preparing for that. We could expect a convergence on the inter-bank market, especially in the interest rates.

Based on that we have set two possible intervals for the expected long-term interest rates. They are established in the following manner; the BRIBOR and PRIBOR overnight rates from 2005 are expected to be in the interval of one standard deviation from the mean values of EURIBOR overnight rates in 2006. We have computed the mean value and the standard deviation of the EURIBOR in 2006 for different months and also for different maturities. The final intervals, summarized in Table 7.1, are:

$$I_\theta = [\theta_d, \theta_u] = [\mu_j - \sigma_j, \mu_j + \sigma_j]$$

where $\mu_j = E(R_j^i)$, $\sigma_j = \sqrt{D(R_j^i)}$, R_j^i -EURIBOR rates, i stands for one month time horizon and $j \in \{on, 1w\}$ is the maturity. In the case of the convergence to the shorter maturity interval, the I_θ is much more wider (see Figure 7.1) than in the case of longer maturities such as the one week case. The reason is the greater volatility of the overnight EURIBOR rates. Of course the I_θ interval could be determined in another way.

Now, the externally added information is given. We can continue with the calibration steps to obtain the interval for the market price of risk. Demonstration of the calibrated intervals I_λ is shown on the Figure 7.2. The calibration has been done in a sample period of time for the BRIBOR and PRIBOR rates in 2005, with the expectation of the convergence to the EURIBOR in 2006 used the CIR model. The first step of the calibration was the same as for the internal calibration method. We have obtained $(\check{\beta}, \check{\xi}, \check{\varrho})$ for any given λ. In the second step we added our expectation about the I_θ and the outcome was the interval for the market price of risk.

As we can see from the results for the market price of risk, $\lambda < 0$ in most of the cases for BRIBOR. The opposite is true for the market price of risk interval for PRIBOR rates. Since λ is a constant determining the risk premium factor $1 - \lambda B(\tau)$ the expected rate of bond return $r^* = (1 - \lambda B(\tau))r$ is greater than the overnight rate r if and only if $\lambda < 0$.

One possible conclusion from such term structure behavior might be that the chosen target interval I_θ for BRIBOR based on the EURIBOR 2006 data verifies greater expected rate of bond return r^* compared to the overnight rate. On the other hand, for the Czech data the target interval I_θ turns to be artficialy pulled up because of positive values of λ leading to lower expected rate of bond return r^* compared to the overnight rate. If we push down the target interval I_θ (i.e. we decrease θ_d and θ_u for which $\lambda_u < 0$) then we achieve greater expected rate of bond return r^* compared to the overnight rate.

Now it is interesting to compare the results of calibration based on the internal and external methodology. In Subsection 6.3.3 we have proposed the outcomes of the internal calibration based on binding of means. As we can see the interval for the market price of risk I_λ is in the most of the cases negative for BRIBOR and PRIBOR rates (see Table 6.3 and 6.5). It means that the predetermined interval of the expected long-term interest rate I_θ is set in a manner proving the condition that the expected rate of bond return is greater than the overnight rate. It does not necessarily mean that the expert could not better adjust his/her expectation on I_θ to obtain the same outcome.

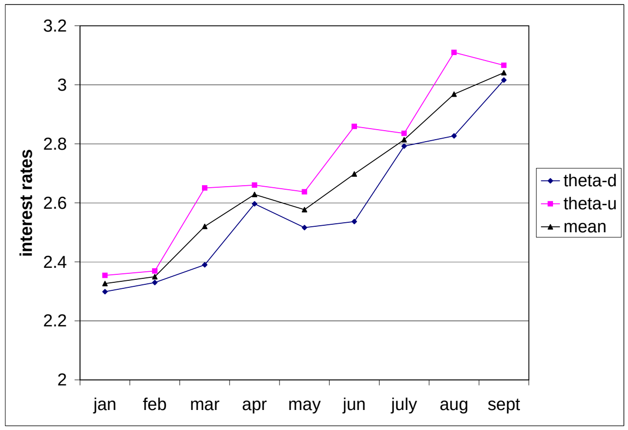


Figure 7.1: Expected long-term interest rate interval $I_\theta = [\theta_d, \theta_u]$, case of overnight EURIBOR rates.

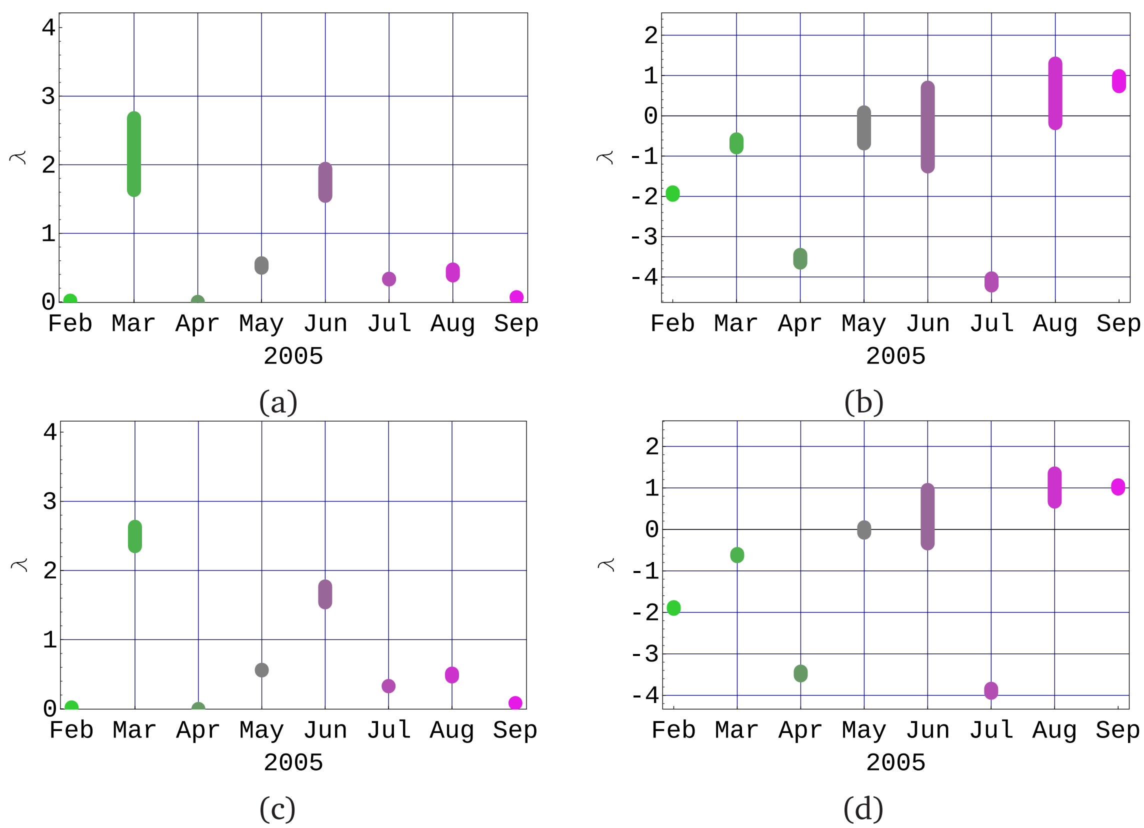


Figure 7.2: Results of calibration based on expected long-term interest rate interval for PRIBOR (left) and BRIBOR (right) rates for different expectations, convergence to overnight (a)-(b) and 1-week (c)-(d) interest rates of EURIBOR.

Chapter 8

Discussion

The main goal of this thesis was to identify and analyze a lot of possibilities of using one factor interest rate models in the context of the European and emerging (new member states) markets. Of course, there exist many different papers dealing with these models, but many of them are focusing only on particular aspect or one issue connected with the one factor models. We attempted to add a new value and insight in this theses, not just a general overview of the existing calibration and estimation techniques and their connection to the interest rate models and derivatives. We proposed new and hopefully interesting issues and topics which came out during the analysis and expertise of the possibilities of usage of these models for real data.

The core results of this theses could be divided into two main parts. The first part is the so-called internal calibration method. This method is based on the new two phase minmax optimization method for parameter estimation of the CIR and Vašiček one factor interest rate model. It is based on minimization of the loss functional together with the maximization of the likelihood function restricted to the set of minimizers. We have tested the estimation method on various term structures including stable west Europe inter-bank offered rates as well as those of the new EU member states. Based on our results of parameter estimation for the CIR one factor model we can state that the western European term structure data are better described with CIR model compared to the new EU member states represented by Central European countries. Our methodology can be applied in order to estimate CIR parameters for EURO-LIBOR, USD-LIBOR and EURIBOR term structures. Interestingly, to some extent, it could be also applied for estimation of CIR parameters for the Czech PRIBOR term structure. On the other hand, we can observe, at least partial, quantitative failure of the CIR model for other Central European term structures.

We measured the quality of non-linear regression by mean of non-linear R^2 ratio.

It is very important to emphasize the possibility of the parameter reduction in the CIR and Vašiček model too. This is utilized during all introduced methods and approaches in this theses. Because of integrated view on interest rate models calibration we sketched a way how to calibrate two factor models including, in particular, the Fong-Vašiček model.

Before turning to the second part of our analysis and results, we insert a somehow new view on the internal calibration method. This is the calibration based on binding interval approach. It means that we do not try to calibrate all the parameters of the one factor models to be a point, but we allow some freedom of a parameter. The endogenity of this method remains, because all the parameters are obtained from the data.

Final part of the theses deals with the external calibration method. In this case the exogenity of this approach is coming from an externally provided parameter. Based on that information which is the expected long-term interest rate interval we have calibrated the remained parameters of the one factor interest rate models. We assumed that the BRIBOR and PRIBOR rates in 2005 will converge to the mean values of EURIBOR with specific maturity in 2006. We prefer the results for BRIBOR from the point of view of negative market price of risk. Our assumption about the expected long-term interest rate interval is not justified for PRIBOR. The plausible expectation would be a lower interest rate interval.

The main contributions of this thesis are:

1. introduction of the new variables for the CIR and Vašiček models to reduce the four dimensional parameter space of the one factor interest rate models into three reduced parameters,

2. transformation of the explicit solution to the PDE for the zero-coupon bond price by means of the new variables,

3. proposition of the loss functional in an aggregated form which is used in the optimization procedure,

4. introduction of a new calibration technique called the minmax procedure consisting of two steps,

5. measuring the quality of the fit by the non-linear R^2 ratio and the maximum likelihood ratio,

6. suggestion of other internal calibration methods (using not only the cummulative statistics of the yield curves) based on mean value binding,

7. proposition of the external calibration methodology utilizing an externally provided information,

8. analysis and comparison of the results of calibration for Central and Western European countries,

9. extension to the multi factor model case.

Results included in points (1)–(5) have been published in author's papers [64, 65]. According to our best knowledge and careful inspection of literature, parts (1)–(9) are new and original contributions of this thesis in the field of parameter calibration of term structure models.

Chapter 9

List of symbols

BRIBOR - Bratislava inter-bank offered rate
CIR - Cox-Ingersoll-Ross
$D(X)$ - dispersion of X
EONIA - Euro overnight index average
ES - Evolution strategies
EU - European Union
EURIBOR - Euro inter-bank offered rate
EURO-LIBOR - London inter-bank offered rate in EUR
$E(X)$ - mean of X
GMM - Generalized method of moments
HJM - Heath-Jarrow-Morton
I_θ - expected long-term interest rate interval
I_λ - market price of risk interval
I_κ - speed of reversion interval
KFA - Kalman filter approach
LMM - LIBOR market model
LF - likelihood function
MCMC - Markov Chain Monte Carlo
MLM - Maximum likelihood method
MLR - Maximum likelihood ratio
P, $P(t,T)$, $P(t,T,r)$, $P(\tau,r)$ - zero-coupon bond price
PRIBOR - Prague inter-bank offered rate
r^* - risk premium factor

r_d^* - lower bound for the risk premium factor
r_u^* - upper bound for the risk premium factor
$R(t,T)$ - yield to maturity
USD-LIBOR - London inter-bank offered rate in USD
U - loss functional
κ - speed of reversion of the short rate process
σ - volatility of the short rate process
θ - expected long-term interest rate of the short rate process
γ - constant for the CIR and Vašiček model specification
λ - market price of risk
τ - time to maturity, $\tau = T - t$
T - maturity of zero-coupon bond
t - actual time

Index

T

U

V

W

Y

Z

Chapter 10

CD ROM Content

- **Thesis**
  thesis.pdf
  A. Urbánová Csajková: Calibration of term structure models

- **Papers**
  Paper-CEJOR
  Paper-CEJOR.pdf
  D. Ševcovic and A. Urbánová Csajková: Calibration of one factor interest rate models, Journal of Electrical Engineering, 55, No. 12/s (2004), 46–50.

  Proceedings-EUBA
  Paper-EUBA.pdf
  A. Urbánová Csajková: Min-max calibration for interest rate models and its application to central European financial markets, Proceedings of 4^{th} International scientific seminar of doctoral students, May 21 (2004), 371–376.

  Proceedings-ISCAM
  Paper-JEE.pdf
  D. Ševcovic and A. Urbánová Csajková: On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model, Central European J. of Operation Research 13 (2005), 169–188.

- **Slides**
  Slides-Algoritmy
  slides.pdf

D. Ševcovic (co-authors: M. Jandacka, A. Urbánová Csajková, B. Stehliková): On some nonlinear problems arising in derivative pricing. Algoritmy 2005, Podbanské

Slides-Wien-Bratislava
slides.pdf
D. Ševcovic and A. Urbánová Csajková: A two phase method for callibration of interest rate models and its application to Central European Interbank Offer Rates. The first Bratislava-Vienna economic meeting, May 7, 2004 Bratislava

- **Source-Codes**
  Data-graphs
  Intervaly-Theta
  Min-Max-Method-Evolution-Strategies
  Minimalizacia-Mathematica
  Rezidualy
  Term-structure-data

Bibliography

[1] Ait-Sahalia, Y., *Testing Continuous-Time Models of the Spot Interest Rate,* Review of Financial Studies **9** (1996), 385–426.

[2] Ait-Sahalia, Y., *Maximum likelihood estimation of discretely sampled diffusions: a closed-form approach,* Econometrica **70** (2002), 223–262.

[3] Algower, E. L., and Georg, K., *Numerical continuation methods,* New York, Heidelberg, Berlin: Springer-Verlag, 1990.

[4] Antol, M., *Analýza jedno a dvojfaktorových modelov derivátov úrokových mier,* Bratislava, Diplomová práca FMFI UK, 2006.

[5] Babbs, S. H. and Nowman, K. B., *An Application of Generalized Vasicek Term Structure Models to the UK Gilt-Edged Market: A Kalman Filtering Analysis,* Applied Financial Economics **8** (1998), 637–644.

[6] Babbs, S. H. and Nowman, K. B., *Kalman filtering of generalized Vasicek term structure models,* Journal of Financial and Quantitative Analysis **34** (1999), 115–130.

[7] Ball, C., and Torous, W., *Unit roots and the estimation of interest rate dynamics,* Journal of Empirical Finance **3** (1996), 215–238.

[8] Ball, C., and Torous, W., *The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence,* Journal of Finance **54** (1999), 2339–2359.

[9] Bergstrom, A. R., *Gaussian estimation of structural parameters in higher order continuous time dynamic models,* Econometrica **51** (1983), 117–152.

[10] Bergstrom, A. R., *Continuous time stochastic models and issues of aggregation over time,* in: Z. Griliches and M.D. Intriligator, Eds., Handbook of Econometrics, Vol II, Elsevier Science Amsterdam, 1984.

[11] Bergstrom, A. R., *The estimation of parameters in nonstationary higher-order continuous time dynamic models,* Econometric Theory **1** (1985), 369–385.

[12] Besag, J., *Spatial interaction and the statistical analysis of lattice systems (with discussion),* Journal of the Royal Statistical Society **36** (1974), 192–236.

[13] Black, F., Derman, E., and Toy, W, *A one-factor model of interest rates and its application to treasury bond options,* Financial Analysts Journal, **46**(1) (1990), 33–39.

[14] Boyle, P. P., Tan, K. S., and Tian, W., *Calibrating the Black-Derman-Toy model: some theoretical results,* Applied Mathematical Finance **8** (2001), 27–48.

[15] Brace, A., Gatarek, D., and Musiela, M., *The Market Model of Interest Rate Dynamics,* Math. Finance **7** (1997), 127–154.

[16] Brandt M.W., and He P., *Simulated likelihood estimation of affine term structure models from panel data,* Working Paper Series (2006). Available at SSRN: `ssrn.com/abstract=885682`

[17] Brown, S. J. and Dybvig, P. H., *The empirical implications of the Cox, Ingersoll, and Ross theory of the term structure of interest rates,* Journal of Finance **41** (1986), 617–632.

[18] Brandt, M. W., and Santa-Clara, P., *Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets,* Journal of Financial Economics **63** (2002), 161–210.

[19] Carlin B. P., Polson N. G., and Stoffer D. S., *A Monte Carlo Approach ton Nonnormal and Nonlinear State-Space Modeling,* JASA **87** (1992), 493–500.

[20] Chan, K. C., Karolyi, G. A., Longstaff, F. A., and Sanders, A. B., *An Empirical Comparison of Alternative Models of the Short-Term Interest Rate,* Journal of Finance **47** (1992), 1209–1227.

[21] Chen, R. R., *Interest Rate Dynamics, Derivatives Pricing, and Risk Management,* Lecture Notes in Econometrics and Mathematical Systems 435, Springer 1996.

[22] Chen, K. C., and Scott, L., *Multi-Factor Cox-Ingersoll-Ross Model of the Term Structure: Estimates and Tests From a Kalman Filter Model,* Journal of Real Estate Finance and Economics **27** (2003), 143–172.

[23] Cortazar, G., Schwartz, E. S., and Naranjo, L., *Term structure estimation in low-frequency transaction markets: A Kalman filter approach with incomplete panel-data, (March 2004)* EFA 2004 Maastricht Meetings Paper No. 3102. Available at SSRN: `ssrn.com/abstract=567090`

[24] De Jong, F., *Time-series and cross-section information in affine term structure models,* Journal of Business and Economic Statistics **18** (2000), 300–314.

[25] De Jong, F., and Santa-Clara, P., *The dynamics of the forward interest rate curve: A formulation with state variables,* Journal of Financial and Quantitative Analysis **34** (1999), 131–157.

[26] De Rossi, G., *Maximum likelihood estimation of the Cox-Ingersoll-Ross model using particle filters,* Computing in Economics and Finance (2004), No 302.

[27] Duan, J.-CH., Gauthier, G., Simonato, J.-G., and Zaanoun, S., *Maximum Likelihood Estimation of Structural Credit Spread Models – Deterministic and Stochastic Interest Rates,* (2002). Available at: `www.globalriskguard.com/resources/deriv/MLECreditSpread.pdf`

[28] Duan, J.-CH., and Simonato, J.-G., *Estimating and testing exponential-affine term structure models by Kalman filter,* Review of Quantitative Finance and Accounting **13** (1999), 102–127.

[29] Duffee, G. R., *Estimating the price of default risk,* Review of Financial Studies **12** (1999), 197–226.

[30] Durham, G. B., and Gallant, A. R., *Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes,* Journal of Business and Economic Statistics **20** (2002), 335–338.

[31] Elerian O., Chib S., and Shephard N., *Likelihood inference for discretely observed non-linear diffusions,* Econometrica **69** (2001), 959–993.

[32] Eraker B., *MCMC analysis of diffusion models with application to finance,* Journal of Business and Economic Statistics **19**(2) (2001), 177–191.

[33] Fong, H. G., and Vasicek, O. A., *Fixed-Income Volatility Management,* Journal of Portfolio Management (1991 Summer), 41–42.

[34] Fruhwirth-Schnatter S., and Geyer, A. L. J., *Bayesian Estimation of Econometric Multi-factor Cox-Ingersoll-Ross-Models of the Term Structure of Interest Rates via MCMC Methods,* Working Paper, Available at: `www.wu-wien.ac.at/usr/or/geyer/papers/mcmc2.pdf`

[35] Geyer, A. L. J., and Pichler, S., *A state-space approach to estimate and test multifactor Cox-Ingersoll-Ross models of the term structure,* Journal of Financial Research **23** (1999), 107–130.

[36] Gibbons, M. R., and Ramaswamy, K., *The term structure of interest rates: empirical evidence,* Columbia - Center for Futures Markets Papers (1988), No 174.

[37] Gibbons, M. R., and Ramaswamy, K., *A test of the Cox, Ingersoll, and Ross model of the term structure,* Review of Financial Studies **6** (1993), 619–658 .

[38] Hansen, L. P., *Large sample properties of Generalized Method of Moments estimator,* Econometrica **50** (1982), 1029–1054.

[39] Harvey, C., *The real term structure and consumption growth,* Journal of Financial Economics **22** (1988), 305–333.

[40] Harvey, A., *Forecasting, structural time series models and the Kalman filter,* Cambridge University Press, Cambridge 1989.

[41] Hornišová, K., *Intristic Linearization of Nonlinear Regression by Principal Components Method,* Acta Math. Univ. Comenianae **73**(2) (2004), 207–216.

[42] Hull, J., *Options, Futures and Other Derivative Securities,* New York: Prentice Hall, 1989.

[43] Jacquier E., Polson N. G., and Rossi P. *Bayesian analysis of stochastic volatility models,* Journal of Business and Economic Statistics **12**(4) (1994), 371–389.

[44] Jamshidian, F., *The one-factor Gaussian interest rate model: Theory and implementation,* Working paper, Merril Lynch Capital Markets, 1988.

[45] Jamshidian, F., *LIBOR and Swap Market Models and Measures,* Finance and Stochastics **1** (1997), 293–330.

[46] Jensen, M. B., *Efficient method of moments estimation of the Longstaff and Schwartz interest rate model,* Working paper series **52** (2000). Available at: `www.cls.dk/caf/wp/wp-52.pdf`

[47] Kwok, Y. K., *Mathematical Models of Financial Derivatives,* New York, Heidelberg, Berlin: Springer-Verlag, 1988.

[48] Lamoureaux, C., and Witte, H. D., *Empirical analysis of the yield curve: the information in the data viewed through the window of Cox, Ingersoll and Ross,* Journal of Finance **57** (2002), 1479–1520.

[49] Li, S., *An Empirical Study of Australian Short-term Interest Rate: A Comparison of Single Factor Models,* School of Finance and Business Economics Working Paper Series (2000). Available at: `www.business.ecu.edu.au/schools/afe/wps/papers/pdfs/wp0011sl.pdf`

[50] Longstaff, F. A., *A nonlinear general equilibrium model of the term structure of interest rates,* Journal of Financial Economics **23** (1989), 195–224.

[51] Lund, J., *Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rate,* Working Paper (1997a). Available at: `www.jesperlund.com/papers/kflin.pdf`

[52] Lund, J., *Non-linear Kalman filtering techniques for term structure models,* Working Paper (1997b). Available at:
`citeseer.ist.psu.edu/cache/papers/cs/3892/http:zSzzSzwww.hha.dk`
`zSz~jelzSzpaperszSzkfnlin1.pdf/lund97nonlinear.pdf`

[53] Miltersen, K. R., Sandmann, K., and Sondermann, D., *Closed form solutions for term structure derivatives with log-normal interest rates,* Journal of Finance **1** (1997), 409–430.

[54] Olšárová, L., *Quantitative comparison of calibration methods for interest rate,* Journal of Electrical Engineering **56**, No.12/s (2005),110–112.

[55] Pázman, A., *Geometry of the Nonlinear Regression with Prior,* Acta Math. Univ. Comenianae **61**(2) (1992), 363–276.

[56] Pearson, N. D., and Sun, T.-S., *Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model,* Journal of Finance **49** (1994), 1279–1304.

[57] Pedersen, A. R., *A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations,* Scandinavian Journal of Statistics **22** (1995), 55–71.

[58] Rechenberg, I., *Evolutionstrategie: Optimierung Technischer Systeme nach Prinziepen der Biologishen Evolution,* Frommann-Holzboog, Stuttgart 1973.

[59] Rendleman, R. and Bartter, B., *The pricing of options on debt securities,* Journal of Financial and Quantitative Analysis, **15** (1980), 11-24.

[60] Santa-Clara, P., *Simulated likelihood estimation of diffusion with an application to the short-term interest rate,* Mimeo, UCLA (1995).

[61] Schwefel, H.-P., *Numerische Optimierung von Computermodelen Mittels der Evolutionsstrategie,* Vol. 26 of Interdisciplinary System Research, Birkhäuser, Basel 1977.

[62] Schwefel, H.-P., *Evolution and Optimum Seeking,* Sixth- Generation Computer Technology Series, Wiley, New York 1995.

[63] Schwefel, H.-P., and Rudolph, G., *Contemporary Evolution Strategies,* In: V F. Morán, A. Moreno, J.J. Merelo, P. Chacón (eds) Advances in Artificial Life, Third International Conference on Artificial Life, Lecture Notes in Artificial Intelligence Vol. 929, Springer Berlin 1998, 893–907.

[64] Ševčovič, D. and Urbánová Csajková, A., *Calibration of one factor interest rate models,* Journal of Electrical Engineering **55**, No. 12/s (2004), 46-50.

[65] Ševčovič, D. and Urbánová Csajková, A., *On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model,* Central European Journal of Operational Research **13** (2005), 169-188.

[66] Selby, M., and Strickland, Ch., *Computing the Fong and Vašíček Pure Discount Bond Price Formula,* Journal of Fixed Income **5** (1995), 78–84.

[67] Stehlíková, B., *Kalibrácia jednofaktorových modelov úrokových mier,* Rigorous thesis, 2005.

[68] Stehlíková, B., *Analýza dvojfaktorového modelu vývoja úrokovej miery so stochastickou volatilitou,* Diploma thesis, 2004.

[69] Stehlíková, B. and Ševčovič, D., *On a volatility averaging in a two factor interest rate model,* Proc. of Algoritmy 2005, 325-333.

[70] Stehlíková, B., *Modeling volatility clusters with application to two-factor interest rate models,* Journal of Electrical Engineering **56**, No. 12/s (2002), 90–93

[71] Stehlíková, B., *Fast mean reverting volatility in Fong-Vasicek model of interest rates,* Journal of Electrical Engineering **57**, No. 7/s (2006), 65–67

[72] Štulajter, F., *Mean Square Error Matrix of an Approximate Least Squares Estimator in a Nonlinear Regression Model with Correlated Errors,* Acta Math. Univ. Comenianae **61**(2) (1992), 251–261.

[73] Takahashi, A. and Sato, S., *A Monte Carlo filtering approach for estimating the term structure of interest rates,* Annuals of the Inst. of Statistical Mathematics **53** (2001), 50–62.

[74] Vojtek, M., *Calibration of Interest Rate Models - Transition Market Case,* Discussion Paper Series, CERGE-EI, 2004.

[75] Wehlau, D., *The Creation of Financial Markets and Pension Reforms in Central and Eastern Europe - Are there any Links?,* EPOC - Conference on ”Privatisation of Public Pension Systems - Forces, Experience, Prospects”, 2003. Available at: `www.epoc.uni-bremen.de/publications/pup2003/files/Vienna_Wehlau_2003.PDF`