FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS
COMENIUS UNIVERSITY BRATISLAVA

Mathematical analysis of term structure models

Dissertation Thesis

Beáta Stehlíková

Supervisor: Daniel Ševčovič

Bratislava 2008

Department of Applied Mathematics and Statistics
Faculty of Mathematics, Physics, and Informatics
Comenius University
Mlynská dolina
842 48 Bratislava
Slovakia
http://www.iam.fmph.uniba.sk/institute/stehlikova

Doctoral Thesis in Applied Mathematics

I would like to thank my supervisor Daniel Ševčovič for his valuable ideas and guidance in my research.

Contents

Chapter 1

Introduction

Term structure models describe a functional dependence between the time to maturity of a discount bond and its present price. Yield of bonds, as a function of maturity, forms a term structure of interest rates. Figure 1.1 shows the different shapes of term structures observed on the market based on data by Bloomberg[1].

Continuous interest rate models are often formulated in terms of stochastic differential equations for the instantaneous interest rate (or short rate) as well as equations for other relevant quantities like e.g. volatility of the short rate process. In one-factor models there is a single stochastic differential equation for the short rate. The volatility of the short rate process is given in a deterministic way. It is assumed to be constant (the Vasicek model) or it is a function of the short rate itself (the Cox, Ingersoll, and Ross model). Beside these two simple models there is a wide range of other models including, in particular, the Chan-Karolyi-Longstaff-Sanders model, the Hull-White model and many others. Based on this assumption made on the form of the short rate process one-can derive a linear scalar parabolic equation for the bond price as function of the current short rate and time to maturity. The reader is referred to the books by Kwok [29] and Brigo and Mercurio [12] for detailed discussion on applications and properties of one-factor interest rate models.

In one-factor models, term structure of interest rates is a function of a short rate and model parameters. However, it means that once the parameters of the model are given, the term structure corresponding to a given short rate is uniquely determined. This is a simplification of the reality, as it can be seen in Figures 1.2 and 1.3, showing

[1] http://www.bloomberg.com/

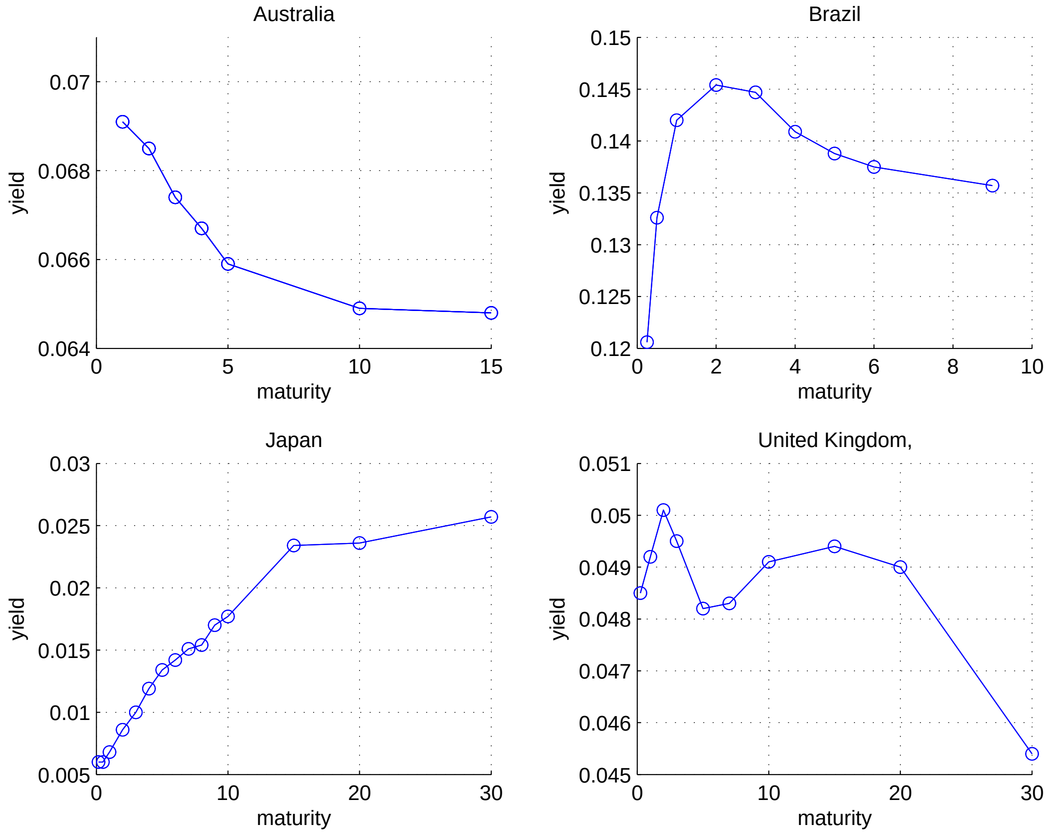

Figure 1.1: Examples of yield curves of governmental bonds: Australia, Brazil, Japan, United Kingdom (27th May 2008).

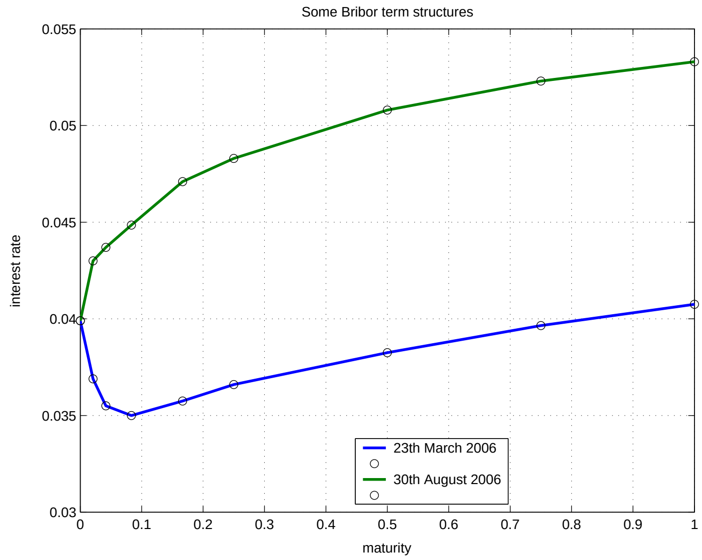

Figure 1.2: Examples of real Bribor term structures.

the examples from Bribor[2] and Euribor[3] data. To capture this feature, two-factor models are introduced. In the two-factor models there are two sources of uncertainty yielding different term structures for the same short rate. They may depend on the value of the other factor. Moreover, two-factor models have more variety of possible shapes of term structures. Again, the reader is referred to the books by Kwok [29] and Brigo and Mercurio [12] for detailed discussion on two-factor interest rate models.

There are several ways of incorporating the second stochastic factor. It is reasonable to conjecture that the market changes the volatility of the underlying process for the short rate. In the so-called two-factor models with a stochastic volatility we allow the volatility to have a stochastic behavior driven by another stochastic differential equation. We focus our attention on the Fong-Vasicek model in which the volatility of the short rate process satisfies a Bessel square root stochastic differential equation. As a consequence of the multidimensional Itō's lemma the corresponding equation for the bond price is a linear parabolic equation in two space dimensions. These spatial dimensions correspond to the short rate and volatility. It is well known that its density distribution of a stochastic process is a solution to the Fokker-Planck partial differential equation and can be expressed analytically in the case the volatility undergoes the Bessel square root process (see e.g. [27]). Knowing the density distribution of the stochastic volatility we are yet able to perform averaging of the bond price and the term structure with respect to volatility. Unlike the short rate which is known from the market data on daily basis, the volatility of the short rate process is unknown. Therefore such a volatility averaging is of special importance

[2] http://www.nbs.sk
[3] http://www.euribor.org/

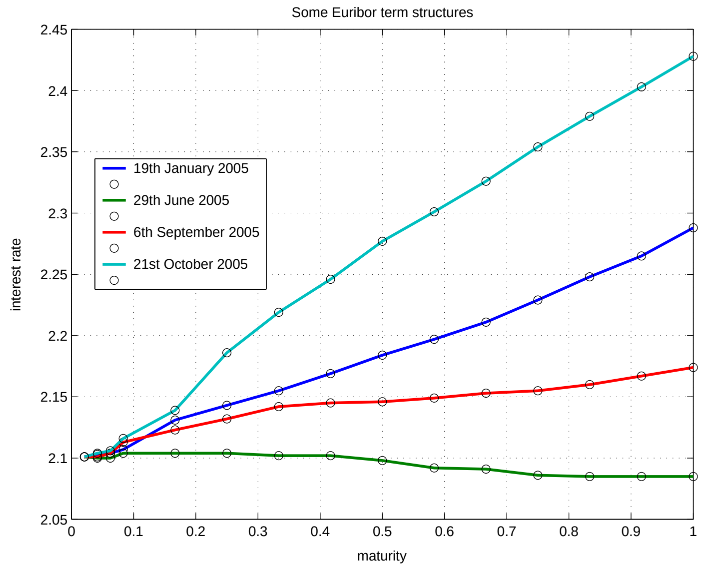

Figure 1.3: Examples of real Euribor term structures.

for practitioners.

Another popular approach in two-factor models is based on describing the short rate as a sum of two components, each of them driven by a stochastic process. These two factors can be interpreted as trend and speculative components. We will study two-factor Vasicek and Cox-Ingersoll-Ross models constructed in this way. The bond prices are again solutions of a linear parabolic equation in two space dimensions where the dimensions correspond to the two factors of the short rate. The density distribution of the factors can be derived and we consider the conditional distribution with respect to the observed level of the short rate. Afterwards, we study averaging of the bond price and the term structure with respect this conditional distribution, as it is only the short rate (and not its two factors) that is observable variable on the market.

The thesis is organized as follows. In the following section we present goals of the thesis in a more detail. In the third section we do a survey of one-factor models. Fourth and fifth sections contain results regarding one-factor models. In fourth section we study the approximate analytical solution for a class of one-factor models for which the closed form bond prices are not available. We summarize results of the recent author's paper [47]. These approximations are used in fifth chapter to calibrate the models. Sixth chapter gives a survey of two-factor models. The next three chapters deal with three particular two-factor models - two-factor Vasicek (chapter 7), two-factor Cox-Ingersoll-Ross (chapter 8) and Fong-Vasicek (chapter 9) models - and the averaged bond prices and term structures with respect to unobservable quantities of the models. Results of these chapters are based on author's papers [42], [45] and [46]. In the case of Fong-Vasicek model we study also the fast mean reverting volatility (according to author's paper [44]). Then we generalize this model to capture volatility clustering, showing the results from [43].

Chapter 2

Goals of the thesis

In the thesis we study and analyze several questions and problems that are related to short rate interest rate models. The main goals of the thesis can be summarized as follows:

1. **Approximate analytical solution for one-factor models.** We study the approximate analytical solution for bond prices derived by Choi and Wirjanto in [17]. We prove the order of accuracy of their formula and present numerical examples. Afterwards, we provide a new approximation of higher order of accuracy.

2. **Calibration of one-factor models.** We use the approximate analytical solution mentioned above to calibrate one-factor models. We use Nowman's Gaussian estimates to estimate the volatility and the comparison of real term structures with theoretical ones to estimate the drift. Here we also study the question of existence of the estimates, i.e. the existence of maximum of likelihood function. Then we consider different weights when comparing the term structures and we see the differences in estimates caused by different criteria used.

3. **Averaging in two-factor models.** We consider the following two-factor models: two-factor Vasicek, two-factor Cox-Ingersoll-Ross and Fong-Vasicek. In all these models, not all of the factors is are observable on the market. We consider their limiting distribution and compute the distribution of bond priced and interest rates. Afterwards, we compute their averaging, i.e. the expected values with respect to limiting distribution of unobservable factors. The averaged bond prices are functions of maturity and short rate. It is the same dependence as in

one-factor models. Hence we study the question, whether there exists a one-factor model, which yields the same bond prices as the averaged values from the two-factor model. In all the models considered, the answer is negative.

Chapter 3

A survey of one-factor short rate models

In this chapter we describe the most widely used one-factor interest rate models. We consider continuous models for the short rate process in the form

$$dr = \mu(t,r)dt + \sigma(t,r)dw, \tag{3.1}$$

where w is a Wiener process. Recall that a stochastic process $\{w(t), t \geq 0\}$ is called a Wiener process if $w(0) = 0$, every increment $w(t + \Delta t) - w(t)$ has the normal distribution $N(0, \Delta t)$, the increments $w(t_n) - w(t_{n-1})$, $w(t_{n-1}) - w(t_{n-2})$, $\ldots$, $w(t_2) - w(t_1)$ for $0 \leq t_1 < \cdots < t_n$ are independent and paths of the process are continuous (see e.g. [29]). Function μ in (3.1) determines the trend in evolution of the short rate, function σ the nature of stochastic fluctuations. The price of a discount bond $P(\tau, r)$ with time to maturity τ when the value of short rate is r, is known to be a solution of the partial differential equation

$$-\frac{\partial P}{\partial \tau} + (\mu(t,r) - \lambda(t,r)\sigma(t,r))\frac{\partial P}{\partial r} + \frac{\sigma^2(t,r)}{2}\frac{\partial^2 P}{\partial r^2} - rP = 0, \tag{3.2}$$
$$P(0,r) = 1.$$

In derivation of the above equation we employ the method used in [29]. It is based on construction of a riskless portfolio and elimination of a possibility of arbitrage. The main mathematical tool in the derivation is the so-called Itō's lemma. According to Itō's lemma, if the stochastic differential equation (SDE) for a process x is given by

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

then the process $f(t,x)$ (where f is a C^2 smooth function) satisfies the SDE

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

(see e.g. [37]).

If we assume that the short rate follows the process (3.1), then the bond price $P(t,r)$ follows - according to Itō's lemma - a stochastic differential equation

$$\begin{aligned} dP &= \left(\frac{\partial P}{\partial t} + \mu\frac{\partial P}{\partial r} + \frac{\sigma^2}{2}\frac{\partial^2 P}{\partial r^2} \right) dt + \sigma\frac{\partial P}{\partial r} dw \\ &= \mu_B(t,r,T)dt + \sigma_B(t,r,T), \end{aligned} \tag{3.3}$$

where $\mu_B(t,r,T)$ and $\sigma_B(t,r,T)$ denote a drift and volatility of a process describing the price of a bond with the maturity T. We construct a portfolio of two bonds with different maturities. The portfolio is constructed from one bond with maturity T_1 and Δ bonds with maturity T_2. Hence its value π is:

$$\pi = P(t,r,T_1) + \Delta P(t,r,T_2) \tag{3.4}$$

and the change $d\pi$ of its value is given by

$$\begin{aligned} d\pi &= dP(t,r,T_1) + \Delta dP(t,r,T_2) \\ &= (\mu_B(t,r,T_1) + \Delta\mu_B(t,r,T_2))\, dt + (\sigma_B(t,r,T_1) + \Delta\sigma_B(t,r,T_2))\, dw. \end{aligned} \tag{3.5}$$

By choosing Δ such that

$$\Delta = -\frac{\sigma_B(t,r,T_1)}{\sigma_B(t,r,T_2)} \tag{3.6}$$

we eliminate the stochastic part in (3.5). We then obtain a riskless portfolio satisfying

$$d\pi = \left(\mu_B(t,r,T_1) - \frac{\sigma_B(t,r,T_1)}{\sigma_B(t,r,T_2)}\mu_B(t,r,T_1) \right) dt.$$

To avoid a possibility of arbitrage, the return on the portfolio π must be equal to the riskless instantaneous interest rate r. Hence

$$\mu_B(t,r,T_1) - \frac{\sigma_B(t,r,T_1)}{\sigma_B(t,r,T_2)}\mu_B(t,r,T_1) = r\pi,$$

from which, after substituting the value of the portfolio π from (3.4) and (3.6), we obtain

$$\mu_B(t,r,T_1) - \frac{\sigma_B(t,r,T_1)}{\sigma_B(t,r,T_2)}\mu_B(t,r,T_1) = r\left(P(t,r,T_1) - \frac{\sigma_B(t,r,T_1)}{\sigma_B(t,r,T_2)}P(t,r,T_2) \right).$$

This equality implies that

$$\frac{\mu_B(t,r,T_1) - rP(t,r,T_1)}{\sigma_B(t,r,T_1)} = \frac{\mu_B(t,r,T_2) - rP(t,r,T_2)}{\sigma_B(t,r,T_2)}.$$

Since maturities T_1 and T_2 were chosen arbitrarily, the above expression has to be independent of the maturity of a bond, i.e. there is a function $\lambda(t,r)$ such that

$$\lambda(t,r) = \frac{\mu_B(t,r,T) - rP(t,r,T)}{\sigma_B(t,r,T)} \tag{3.7}$$

for every maturity T. It is called a market price of risk, as it provides an expected rise of the bond return for the unit rise of risk (see e.g. [29]). Substituting functions μ_B, σ_B into (3.7) yields a partial differential equation (3.2) for bond prices.

Bond prices determine interest rates $R(\tau,r)$ by the formula $P = e^{-R\tau}$, i.e.

$$R(\tau,r) = -\frac{1}{\tau}\log P(\tau,r).$$

One of the first models of the class (3.1) was the Vasicek model [51]. The short rate process follows a stochastic differential equation

$$dr = \kappa(\theta - r)dt + \sigma dw.$$

Deterministic part of the process $\kappa(\theta - r)$ defines a mean reversion process with a limit θ. The speed of reversion is given by the parameter $\kappa > 0$. In this model, for a constant market price of risk λ, the PDE for bond prices (3.2) has an explicit solution of the form

$$P(\tau,r) = A(\tau)e^{-B(\tau)r}. \tag{3.8}$$

Substituting it into equation (3.2) we obtain a system of ordinary differential equations

$$\begin{aligned} \frac{dA(\tau)}{d\tau} &= (\lambda\sigma - \kappa\theta)A(\tau)B(\tau) + \frac{1}{2}\sigma^2 A(\tau)B^2(\tau), \\ \frac{dB(\tau)}{d\tau} &= -\kappa B(\tau) + 1, \end{aligned}$$

and initial conditions $A(0) = 1$, $B(0) = 0$. A solution A, B is given by

$$B(\tau) = \frac{1 - e^{-\kappa\tau}}{\kappa},\; A(\tau) = \exp\left[(B(\tau) - \tau)\left(\theta - \frac{\sigma^2}{2\kappa^2} - \frac{\sigma\lambda}{\kappa}\right) - \frac{\sigma^2 B(\tau)^2}{4\kappa}\right]. \tag{3.9}$$

For a fixed r, interest rates $R(\tau,r)$ are functions of the maturity τ. As $\tau \to \infty$, they converge to the value

$$R_\infty := \lim_{\tau\to\infty} R(\tau,r) = \theta - \frac{\sigma\lambda}{\kappa} - \frac{\sigma^2}{2\kappa^2}, \tag{3.10}$$

which does not depend on r. Depending on the relation between the beginning of the short rate r and the limit R_∞, the term structure has one of the following shapes: it is increasing for $r \leq R_\infty - \frac{\sigma^2}{4\kappa^2}$, decreasing for $r \geq R_\infty + \frac{\sigma^2}{2\kappa^2}$ or having a hump (firstly increasing and then decreasing) for r from the interval $(R_\infty - \frac{\sigma^2}{4\kappa^2}, R_\infty + \frac{\sigma^2}{2\kappa^2})$.

A disadvantage of the Vasicek model consists in its constant volatility (it does not depend on the short rate level) and possibility of negative values of the short rate. These problems have been overcome by the short rate model due to Cox, Ingersoll and Ross [19], in which we assume

$$dr = \kappa(\theta - r)dt + \sigma\sqrt{r}dw.$$

If this process starts from a positive value, then its values in future times are non-negative. Moreover, if $2\kappa\theta \geq \sigma^2$, then these values are positive almost surely (see [19]). Similarly as for the Vasicek, the CIR model also admits an explicit solution to equation (3.2). If the market price of risk $\lambda(t, r)$ equals $\lambda\sqrt{r}$, where λ is a constant, then the solution has a form (3.8), where functions A and B satisfy the system of ODEs:

$$\begin{aligned}\frac{dA(\tau)}{d\tau} &= -\kappa\theta A(\tau)B(\tau),\\ \frac{dB(\tau)}{d\tau} &= -(\kappa + \lambda\sigma)B(\tau) - \frac{1}{2}\sigma^2 B(\tau)^2 + 1,\end{aligned}$$

and initial conditions $A(0) = 1$, $B(0) = 0$. A solution is given by

$$A(\tau) = \left[\frac{2\xi e^{(\xi+\psi)\tau/2}}{(\xi + \psi)(e^{\xi\tau} - 1) + 2\xi}\right]^{2\kappa\theta/\sigma^2}, \; B(\tau) = \frac{2(e^{\xi\tau} - 1)}{(\xi + \psi)(e^{\xi\tau} - 1) + 2\xi}, \tag{3.11}$$

where

$$\psi = \kappa + \lambda\sigma, \; \xi = \sqrt{\psi^2 + 2\sigma^2}.$$

The limit of term structures for $\tau \to \infty$ is $\frac{2\kappa}{\xi+\psi}\theta$ (see [29]).

There are several other models, where the SDE for short rate is given by

$$dr = (a + br)dt + \sigma r^\gamma dw. \tag{3.12}$$

Comparison of these models is a topic of the paper by Chan, Karolyi, Longstaff and Sanders [15]. Using generalized method of moments they estimated the model (3.12) and they studied restrictions on parameters imposed on this models. Their result that the optimal value of the parameter γ is approximately $3/2$ (which is more than previous models assumed), started a broad discussion on the correct form of volatility. Let us note that their result is not universal, e.g. in [3], using the same estimation methodology but for LIBOR rates, γ was estimated to be less than unity (which means that volatility is less than proportional to short rate, unlike in the result due to Chan, Karolyi, Longstaff and Sanders). A modification of generalized method of moments (so called robust generalized method of moments), which is robust to a presence of outliers, was developed in [22].

Another popular method for parameter estimation are Nowmans' Gaussian estimates [32], based on approximating the likelihood function. They were used in [23] for a wide range of interest rate markets. For a similar sample of countries, in [50] authors studied the robustness of the estimates with respect to a sample period and use of interest rates with different maturities as short rate proxies. Similarly, different kinds of interest rates as short rate proxies were used in [33] for the Japanese financial market. Estimations for UK and USA data and subsequent study of forecasting power of the models were performed in [14].

There are several other methods based on short rate process, such as quasi maximum likelihood, maximum likelihood based on series expansion of likelihood function by Ait-Sahalia [7], Bayesian methods and others.

Term structure can be computed analytically in the case of Vasicek and CIR models. This feature was used in the calibration methodologies discussed in [48] and [49]. They are based on the minimization of weighted squares of differences between real and theoretical interest rates, followed by optimizing Nowman's likelihood function. For general term structure models, numerical techniques for solving the partial differential equation for bond prices can be used. Frequently used method is the Crank-Nicolson scheme, see e.g. [18]. In [41] and its earlier preprint version [9], the Box method for computation of bond and contingent claim prices was introduced. It was compared with the Crank-Nicolson scheme in specific cases, where analytical prices are available and the new scheme led to more accurate results. This method was used to compute interest rates and derivatives prices in [34] (for Japan), [35] (for UK and USA), [36] (for Canada, Hong Kong and USA). Recently, an analytical approximation formula for bond prices was suggested by Choi and Wirjanto in [17] and by Stehlíková and Ševčovič in [47].

As we have already mentioned, in short rate models, for a given process driving the short rate r and market price of risk λ, we obtain a term structure $R(t, T, r)$. Hence we might be interested in models, where the term structure generated in this way, is an accurate approximation of the real one. This can be obtained, if functions μ and σ in (3.1) depend on time t. It leads to a class of models which are referred to as no-arbitrage models.

In Table 3.1 (time independent drift and volatility) and Table 3.2 (time dependent drift and volatility) we present an overview of selected one-factor short rate models and their characteristics.

Table 3.1: One-factor short rate models with time independent drift and volatility (c.f. [12], Table 3.1).

Model	SDE for short rate r	Distribution of r	Analytic formula for bond price
Vasicek	$dr_t = k(\theta - r_t)dt + \sigma dw_t$	normal	yes
CIR	$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r}dw_t$	noncentral χ^2	yes
Dothan	$dr_t = ar_t dt + \sigma r dw_t$	lognormal	yes
Exponential Vasicek	$dr_t = r_t(\eta - a\ln r_t)dt + \sigma r_t dw$	lognormal	no

Table 3.2: No arbitrage one-factor models short rate with time dependent drift and volatility (c.f. [12], Table 3.1).

Model	SDE for short rate r	Distribution of r	Analytic formula for bond price
Hull-White	$dr_t = k(\theta_t - r_t)dt + \sigma dw_t$	normal	yes
Black-Karasinski	$dr_t = r_t(\eta_t - a\ln r_t)dt + \sigma r_t dw_t$	lognormal	no
Mercurio-Moraleda	$dr_t = r_t\left(\eta_t - \left(\lambda - \frac{\gamma}{1+\gamma t}\right)\ln r_t\right)dt + \sigma r_t dw_t$	lognormal	no

Chapter 4

Approximate analytical solution for a class of one-factor models

In this chapter we discuss the approximate analytical solution for bond prices derived by Choi and Wirjanto in [17]. Authors considered a class of one-factor models and proposed an approximate formula for prices of bonds. In this chapter we derive an accuracy of their approximation and give a new approximation of a higher order of accuracy. Results presented in forthcoming sections 3.1 - 3.4 are included in the recent paper [47] by Stehlíková and Ševčovič. This approximate analytical solution will be used in calibration in section 5 of the thesis.

Models for short rate considered by Choi and Wirjanto in [17] have a form

$$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dw \tag{4.1}$$

under the risk-neutral measure. It corresponds to the real measure process:

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

where $\lambda(t, r)$ is the so called market price of risk. Let us recall that for a general market price of risk function $\lambda(t, r)$, the price P of a zero-coupon bond can be obtained from a solution to the following partial differential equation:

$$-\frac{\partial P}{\partial \tau} + \frac{1}{2}\sigma^2 r^{2\gamma}\frac{\partial^2 P}{\partial r^2} + (\alpha + \beta r)\frac{\partial P}{\partial r} - rP = 0,\ r > 0,\ \tau \in (0, T), \tag{4.2}$$

satisfying the initial condition $P(0, r) = 1$ for all $r > 0$. In what follows, we use the notation $\partial_\tau P$ for $\partial P/\partial \tau$, similarly $\partial_r P$ for $\partial P/\partial r$ and $\partial_r^2 P$ for $\partial^2 P/\partial r^2$.

The main result of the paper [17] is the following approximation P^{ap} for the exact solution P^{ex}:

Theorem 1. *[17, Theorem 2] The approximate analytical solution P^{ap} is given by*

$$\begin{aligned}\ln P^{ap}(\tau,r) &= -rB + \frac{\alpha}{\beta}(\tau - B) + \left(r^{2\gamma} + q\tau\right)\frac{\sigma^2}{4\beta}\left[B^2 + \frac{2}{\beta}(\tau - B)\right] \\ &\quad -q\frac{\sigma^2}{8\beta^2}\left[B^2(2\beta\tau - 1) - 2B\left(2\tau - \frac{3}{\beta}\right) + 2\tau^2 - \frac{6\tau}{\beta}\right] \end{aligned} \tag{4.3}$$

where

$$q(r) = \gamma(2\gamma-1)\sigma^2 r^{2(2\gamma-1)} + 2\gamma r^{2\gamma-1}(\alpha + \beta r) \tag{4.4}$$

and

$$B(\tau) = (e^{\beta\tau} - 1)/\beta. \tag{4.5}$$

Derivation of the formula (4.3) is based on calculating the price as an expected value under a risk neutral measure. The tree property of conditional expectation was used and the integral appearing in the exact price was approximated to obtain a closed form approximation. The reader is referred to [17] for more details of derivation of (4.3).

Authors furthermore showed that such an approximation coincides with the exact solution in the case of the Vasicek model [51]. Moreover, they compared the above approximation with the exact solution of the CIR model which is also known in a closed form (c.f. [19]). Graphical and tabular descriptions of the relative error in the bond prices have been also provided in [17].

Our goal is to derive the order of accuracy of the approximation formula (4.3) by estimating the difference $\ln P^{ap} - \ln P^{ex}$ of logarithms of approximative and exact solutions of the bond valuation equation (4.2). Then, we give a new approximation formula of the higher order and we analyze its order of convergence analytically and numerically.

4.1 Uniqueness of solution to PDE for bond prices

It is worth to note that comparison of approximate and exact solutions is meaningful only if the uniqueness of the exact solution can be guaranteed. The next theorem gives us the uniqueness of a solution to (4.2) satisfying Definition 1 introduced in [47].

Definition 1. *[47, Definition 1] By a complete solution to (4.2) we mean a function $P = P(\tau,r)$ having continuous partial derivatives $\partial_\tau P$, $\partial_r P$, $\partial_r^2 P$ on $Q_T = [0,\infty) \times (0,T)$, satisfying equation (4.2) on Q_T, the initial condition $P(0,r) = 1$ for $r \in [0,\infty)$ and fulfilling the following growth conditions: $|P(\tau,r)| \le Me^{-mr^\delta}$ and $|P_r(\tau,r)| \le M$ for any $r > 0, t \in (0,T)$, where $M, m, \delta > 0$ are constants.*

Theorem 2. *Assume $\frac{1}{2} < \gamma < \frac{3}{2}$ or $\gamma = \frac{1}{2}$ and $2\alpha \geq \sigma^2$. Then there exists a unique complete solution to (4.2).*

Proof: Our aim is to prove that the inequality

$$\frac{d}{d\tau}\int_0^\infty r^\omega P^2 dr \leq K\int_0^\infty r^\omega P^2 dr \tag{4.6}$$

is satisfied by any solution of (4.2) with some constants K and $\omega \geq 0$. It implies the uniqueness of a solution to the PDE (4.2). Indeed, if P_1 and P_2 are two complete solutions of (4.2) with the same initial condition $P(0,r) = 1$, then $P = P_1 - P_2$ is also a solution to (4.2) with $P(0,r) = 0$. Let us define a function

$$y(\tau) = \int_0^\infty r^\omega P^2(\tau, r)dr.$$

Then the inequality (4.6) means $\frac{dy(\tau)}{d\tau} \leq Ky(\tau)$ for $\tau > 0$. It implies:

$$\frac{d}{d\tau}\left(e^{-K\tau}y(\tau)\right) = -Ke^{-K\tau}y(\tau) + e^{-K\tau}\frac{dy(\tau)}{d\tau} \leq 0.$$

Since $y(0) = 0$ and $y(\tau) \geq 0$, it follows that $y(\tau) = 0$ for all τ. Therefore $P(\tau, r) = 0$ for all $\tau \geq 0$, $r \geq 0$ and hence $P_1 \equiv P_2$, as claimed.

Now let us derive inequality (4.6). Multiplying the equation by $r^\omega P$, where $\omega > 0$ and $2\gamma + \omega - 1 > 0$ using the identity $\frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^\omega P^2 dr = \int_0^\infty r^\omega P\partial_\tau P dr$, and integrating with respect to r from 0 to infinity we obtain[1]

$$\frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^\omega P^2 = \frac{\sigma^2}{2}\int_0^\infty r^{2\gamma+\omega}\partial_r^2 PP + \int_0^\infty (\alpha + \beta r)r^\omega \partial_r PP - \int_0^\infty r^{\omega+1}P^2. \tag{4.7}$$

Firstly, we use integration by parts for the following integrals from the above equation:

$$\begin{aligned}\int_0^\infty r^{2\gamma+\omega}\partial_r^2 PP &= -(2\gamma+\omega)\int_0^\infty r^{2\gamma+\omega-1}P\partial_r P - \int_0^\infty r^{2\gamma+\omega}(\partial_r P)^2 \\ &= \frac{1}{2}(2\gamma+\omega)(2\gamma+\omega-1)\int_0^\infty r^{2\gamma+\omega-2}P^2 - \int_0^\infty r^{2\gamma+\omega}(\partial_r P)^2\end{aligned}$$

where we have used the identity $\int_0^\infty r^{\omega+\xi}\partial_r PP = -\frac{\omega+\xi}{2}\int_0^\infty r^{\omega+\xi-1}P^2$ valid for any $\omega, \xi \geq 0$, $\omega + \xi > 0$, and a function P satisfying the decay estimates from Definition 1. Substituting this to (4.7), we end up with the identity

$$\begin{aligned}\frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^\omega P^2 &= \frac{\sigma^2}{4}(2\gamma+\omega)(2\gamma+\omega-1)\int_0^\infty r^{2\gamma+\omega-2}P^2 - \frac{\sigma^2}{2}\int_0^\infty r^{2\gamma+\omega}(\partial_r P)^2 \\ &- \frac{\alpha\omega}{2}\int_0^\infty r^{\omega-1}P^2 - \frac{(\omega+1)\beta}{2}\int_0^\infty r^\omega P^2 - \int_0^\infty r^{\omega+1}P^2. \end{aligned}\tag{4.8}$$

[1] In the sequel, we shall omit the differential dr from the notation

Case 1: $\gamma = \frac{1}{2}$ and $2\alpha \geq \sigma^2$. In the case of CIR model ($\gamma = \frac{1}{2}$) we recall that the condition $2\alpha \geq \sigma^2$ is very well understood as it almost surely guarantees the strict positivity of the stochastic processes r_t, $t \geq 0$, satisfying the stochastic differential equation: $dr = (\alpha + \beta r)\, dt + \sigma\sqrt{r} dw$ (see e.g. [29]).

Subcase 1a: $2\alpha > \sigma^2$. We use the identity (4.8) with $\gamma = 1/2$ and $\omega = \frac{2\alpha}{\sigma^2} - 1 > 0$ to obtain the desired inequality (4.6) with $K = (\omega + 1)\beta$.

Subcase 1b: $2\alpha = \sigma^2$. Using the identity (4.8) with $\omega = 0$ (or simply by multiplying the PDE with P and integrating over $(0, \infty)$) we obtain the inequality (4.6) with $K = \beta$.

Case 2: $\gamma \in \left(\frac{1}{2}, 1\right)$. We use equation (4.7) with $\omega = 2$ and estimate the integral $\int_0^\infty r^{2\gamma} P^2$ by using Hölder's inequality:

$$\int_0^\infty r^{2\gamma} P^2 = \int_0^\infty \left(r^{4\gamma-2} P^{4\gamma-2}\right)\left(r^{2-2\gamma} P^{4-4\gamma}\right) \leq \left(\int_0^\infty r^2 P^2\right)^{2\gamma-1} \left(\int_0^\infty r P^2\right)^{2-2\gamma}.$$

Now it follows from the Young's inequality $ab \leq \frac{1}{p\varepsilon^p} a^p + \frac{1}{q}\varepsilon^q b^q$ valid for $p, q \geq 1$, $\frac{1}{p} + \frac{1}{q} = 1$, that for any $\varepsilon > 0$ we obtain

$$\int_0^\infty r^{2\gamma} P^2 \leq (2\gamma - 1)\left(\frac{1}{\varepsilon}\right)^{\frac{1}{2\gamma-1}} \int_0^\infty r^2 P^2 + (2 - 2\gamma)\varepsilon^{\frac{1}{2\gamma-2}} \int_0^\infty r P^2.$$

Again using (4.8) with $\omega = 2$ and the above estimate we obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^2 P^2 &\leq \frac{\sigma^2}{2}(\gamma+1)(2\gamma+1)\int_0^\infty r^{2\gamma} P^2 - \alpha \int_0^\infty r P^2 - \frac{3\beta}{2}\int_0^\infty r^2 P^2 \\ &\leq K \int_0^\infty r^2 P^2 + \left(\sigma^2(\gamma+1)(2\gamma+1)(1-\gamma)\varepsilon^{\frac{1}{2-2\gamma}} - \alpha\right)\int_0^\infty r P^2. \end{aligned}$$

where $K = \frac{\sigma^2}{2}(\gamma+1)(2\gamma+1)(2\gamma-1)\left(\frac{1}{\varepsilon}\right)^{\frac{1}{2\gamma-1}} - \frac{3\beta}{2}$. By choosing $\varepsilon > 0$ sufficiently small such that $\sigma^2(\gamma+1)(2\gamma+1)(1-\gamma)\varepsilon^{\frac{1}{2-2\gamma}} - \alpha < 0$, we finally obtain the desired inequality $\frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^2 P^2 \leq K \int_0^\infty r^2 P^2$.

Case 3: $\gamma = 1$. We again use equation (4.8) with $\omega = 2$. We obtain (4.6) with $K = 3(2\sigma^2 - \beta)$.

Case 4: $\gamma \in \left(1, \frac{3}{2}\right)$. Similarly as in the case $\frac{1}{2} < \gamma < 1$ we make use of the Hölder inequality. We obtain:

$$\int_0^\infty r^{2\gamma} P^2 = \int_0^\infty \left(r^{6-4\gamma} P^{6-4\gamma}\right)\left(r^{6\gamma-6} P^{4\gamma-4}\right) \leq \left(\int_0^\infty r^2 P^2\right)^{3-2\gamma} \left(\int_0^\infty r^3 P^2\right)^{2\gamma-2}$$

and, by Young's inequality, we have, for any $\varepsilon > 0$,

$$\int_0^\infty r^{2\gamma} P^2 \leq (3 - 2\gamma)\left(\frac{1}{\varepsilon}\right)^{\frac{1}{3-2\gamma}} \int_0^\infty r^2 P^2 + (2\gamma - 2)\varepsilon^{\frac{1}{2\gamma-2}} \int_0^\infty r^3 P^2.$$

By (4.8) with $\omega = 2$ we have

$$\begin{aligned}\frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^2P^2 &\leq \frac{\sigma^2}{2}(\gamma+1)(2\gamma+1)\int_0^\infty r^{2\gamma}P^2 - \frac{3\beta}{2}\int_0^\infty r^2P^2 - \int_0^\infty r^3P^2 \\ &\leq K\int_0^\infty r^2P^2 + \left(\sigma^2(\gamma+1)(2\gamma+1)(\gamma-1)\varepsilon^{\frac{1}{2\gamma-2}} - 1\right)\int_0^\infty r^3P^2.\end{aligned}$$

where $K = \frac{\sigma^2}{2}(\gamma+1)(2\gamma+1)(3-2\gamma)\left(\frac{1}{\varepsilon}\right)^{\frac{1}{3-2\gamma}} - \frac{3\beta}{2}$. By choosing $\varepsilon > 0$ sufficiently small such that $\sigma^2(\gamma+1)(2\gamma+1)(\gamma-1)\varepsilon^{\frac{1}{2\gamma-2}} - 1 < 0$ we end up with the desired inequality $\frac{1}{2}\frac{d}{d\tau}\int_0^\infty r^2P^2 \leq K\int_0^\infty r^2P^2$. ♢

4.2 Error estimates for the approximate analytical solution

In this part we derive the order of accuracy for the approximation proposed by Choi and Wirjanto in [17].

Theorem 3. *[47, Theorem 3] Let P^{ap} be the approximative solution given by (4.3) and P^{ex} be the exact bond price given as a unique complete solution to (4.2). Then*

$$\ln P^{ap}(\tau,r) - \ln P^{ex}(\tau,r) = c_5(r)\tau^5 + o(\tau^5)$$

as $\tau \to 0^+$ where

$$\begin{aligned}c_5(r) = &-\frac{1}{120}\gamma r^{2(\gamma-2)}\sigma^2\left[2\alpha^2(-1+2\gamma)r^2 + 4\beta^2\gamma r^4 - 8r^{3+2\gamma}\sigma^2\right. \\ &+2\beta(1-5\gamma+6\gamma^2)r^{2(1+\gamma)}\sigma^2 + \sigma^4r^{4\gamma}(2\gamma-1)^2(4\gamma-3) \\ &\left.+2\alpha r\left(\beta(-1+4\gamma)r^2 + (2\gamma-1)(3\gamma-2)r^{2\gamma}\sigma^2\right)\right].\end{aligned} \tag{4.9}$$

Convergence is uniform w. r. to r on compact subintervals $[r_1, r_2] \subset (0,\infty)$.

Proof: Recall that the exact bond price $P^{ex}(\tau,r)$ for the model (4.1) is given by a solution of PDE (4.2). Let us define the following auxiliary function: $f^{ex}(\tau,r) = \ln P^{ex}(\tau,r)$. Clearly, $\partial_\tau P^{ex} = P^{ex}\partial_\tau f^{ex}, \partial_r P^{ex} = P^{ex}\partial_r f^{ex}$ and $\partial_r^2 P^{ex} = P^{ex}\left[(\partial_r f^{ex})^2 + \partial_r^2 f^{ex}\right]$. Hence the PDE for the function f^{ex} reads as follows:

$$-\partial_\tau f^{ex} + \frac{1}{2}\sigma^2 r^{2\gamma}\left[(\partial_r f^{ex})^2 + \partial_r^2 f^{ex}\right] + (\alpha+\beta r)\partial_r f^{ex} - r = 0. \tag{4.10}$$

Substitution of $f^{ap} = \ln P^{ap}$ into equation (4.10) yields a nontrivial right-hand side $h(\tau,r)$ for the equation for the approximative solution f^{ap}:

$$-\partial_\tau f^{ap} + \frac{1}{2}\sigma^2 r^{2\gamma}\left[(\partial_r f^{ap})^2 + \partial_r^2 f^{ap}\right] + (\alpha+\beta r)\partial_r f^{ap} - r = h(\tau,r). \tag{4.11}$$

If we insert the approximate solution into (4.2) then, after long but straightforward calculations based on expansion of all terms into a Taylor series expansion in τ we obtain:

$$h(\tau,r) = k_4(r)\tau^4 + k_5(r)\tau^5 + o(\tau^5) \tag{4.12}$$

where the functions k_4 and k_5 are given by

$$\begin{aligned} k_4(r) &= \frac{1}{24}\gamma r^{2(\gamma-2)}\sigma^2 \left[2\alpha^2(-1+2\gamma)r^2 + 4\beta^2\gamma r^4 - 8r^{3+2\gamma}\sigma^2 \right.\\ &\quad +2\beta(1-5\gamma+6\gamma^2)r^{2(1+\gamma)}\sigma^2 + \sigma^4 r^{4\gamma}(-3+16\gamma-28\gamma^2+16\gamma^3)\\ &\quad \left. +2\alpha r\left(\beta(-1+4\gamma)r^2 + (2-7\gamma+6\gamma^2)r^{2\gamma}\sigma^2\right)\right], \end{aligned} \tag{4.13}$$

$$\begin{aligned} k_5(r) &= \frac{\gamma\sigma^2}{120} r^{2(-2+\gamma)} \left[6\alpha^2\beta\left(-1+2\gamma\right)r^2 + 12\beta^3\gamma r^4 - 10(1-2\gamma)^2 r^{1+4\gamma}\sigma^4 \right.\\ &\quad +6\beta^2\sigma^2\left(1-5\gamma+6\gamma^2\right)r^{2(1+\gamma)}\\ &\quad +\beta r^{2\gamma}\sigma^2\left(-10\left(5+2\gamma\right)r^3 + 3(1-2\gamma)^2\left(-3+4\gamma\right)r^{2\gamma}\sigma^2\right)\\ &\quad +2\alpha r\Bigg(3\beta^2\left(-1+4\gamma\right)r^2 + 3\beta\left(2-7\gamma+6\gamma^2\right)r^{2\gamma}\sigma^2\\ &\qquad \left.-5\left(-1+2\gamma\right)r^{1+2\gamma}\sigma^2\Bigg)\right]. \end{aligned} \tag{4.14}$$

Let us consider a function $g(\tau,r) = f^{ap} - f^{ex}$. As $(\partial_r g)^2 = (\partial_r f^{ap})^2 - (\partial_r f^{ex})^2 - 2\partial_r f^{ex}\partial_r g$ we have

$$\begin{aligned} -\partial_\tau g &+ \frac{1}{2}\sigma^2 r^{2\gamma}\left[(\partial_r g)^2 + \left(\partial_r^2 g\right)\right] + (\alpha+\beta r)\partial_r g\\ &= \left\{-\partial_\tau f^{ap} + \frac{1}{2}\sigma^2 r^{2\gamma}\left[(\partial_r f^{ap})^2 + \partial_r^2 f^{ap}\right] + (\alpha+\beta r)\partial_r f^{ap}\right\}\\ &\quad -\left\{-\partial_\tau f^{ex} + \frac{1}{2}\sigma^2 r^{2\gamma}\left[(\partial_r f^{ex})^2 + \left(\partial_r^2 f^{ex}\right)\right] + (\alpha+\beta r)\partial_r f^{ex}\right\}\\ &\quad -\sigma^2 r^{2\gamma}\partial_r f^{ex}\partial_r g\,. \end{aligned}$$

It follows from (4.10) and (4.11) that the function g satisfies the following PDE:

$$-\partial_\tau g + \frac{1}{2}\sigma^2 r^{2\gamma}\left[(\partial_r g)^2 + \partial_r^2 g\right] + (\alpha+\beta r)\partial_r g = h(\tau,r) - \sigma^2 r^{2\gamma}(\partial_r f^{ex})(\partial_r g), \tag{4.15}$$

where $h(\tau,r)$ satisfies (4.12). Let us expand the solution of (4.15) into a Taylor series with respect to τ with coefficients depending on r. We obtain $g(\tau,r) = \sum_{i=0}^\infty c_i(r)\tau^i = \sum_{i=\omega}^\infty c_i(r)\tau^i$, i.e. the first nonzero term in the expansion is $c_\omega(r)\tau^\omega$. Then $\partial_\tau g = \omega c_\omega(r)\tau^{\omega-1} + o(\tau^{\omega-1})$ and $h(\tau,r) = k_4(r)\tau^4 + o(\tau^4)$ as $\tau\to 0^+$. Here the term $k_4(r)$ is given by (4.13). The remaining terms in (4.12) are of the order $o(\tau^{\omega-1})$ as $\tau\to 0^+$. Hence $-\omega c_\omega(\tau) = k_4(r)\tau^4$ from which we deduce $\omega = 5$ and $c_5(r) = -\frac{1}{5}k_4(r)$. It means that $g(\tau,r) = \ln P^{ap}(\tau,r) - \ln P^{ex}(\tau,r) = -\frac{1}{5}k_4(r)\tau^5 + o(\tau^5)$ which completes the proof. $\diamondsuit$

Remark 1. *The function $c_5(r)$ remains bounded as $r \to 0^+$ for the case of the CIR model in which $\gamma = 1/2$ or for the case when $\gamma \geq 1$. More precisely, $\lim_{r\to 0^+} c_5(r) = -\frac{\sigma^2}{120}$ for $\gamma = 1/2$. On the other hand, if $1/2 < \gamma < 1$, then $c_5(r)$ becomes singular, $c_5(r) = O(r^{2(\gamma-1)})$ as $r \to 0^+$.*

Corollary 1. *It follows from Theorem 3 that*

1. *the error in yield curves can be expressed as*

$$R^{ap}(\tau, r) - R^{ex}(\tau, r) = -c_5(r)\tau^4 + o(\tau^4) \text{ as } \tau \to 0^+;$$

2. *the relative error[2] of P is given by*

$$\frac{P^{ap}(\tau, r) - P^{ex}(\tau, r)}{P^{ex}(\tau, r)} = -c_5(r)\tau^5 + o(\tau^5) \text{ as } \tau \to 0^+.$$

Convergence is uniform w. r. to r on compact subintervals $[r_1, r_2] \subset (0, \infty)$.

Proof: The first corollary follows from the formula $R(\tau, r) = -\frac{\ln P(\tau,r)}{\tau}$ for calculating yield curves. To prove the second statement we note that Theorem 3 gives $\ln P^{ap} - \ln P^{ex} = c_5(r)\tau^5 + o(\tau^5)$. Hence $P^{ap}/P^{ex} = e^{c_5(r)\tau^5 + o(\tau^5)} = 1 + c_5(r)\tau^5 + o(\tau^5)$ and therefore $\frac{P^{ap}-P^{ex}}{P^{ex}} = -c_5(r)\tau^5 + o(\tau^5)$. ♢

Remark 2. *For the CIR model with $\gamma = 1/2$ the term $k_4(r)$ defined in (4.13) can be simplified to $\frac{1}{24}\sigma^2 [\alpha\beta + r(\beta^2 - 4\sigma^2)]$ and hence*

$$\ln P^{ap}_{CIR}(\tau, r) - \ln P^{ex}_{CIR}(\tau, r) = -\frac{1}{120}\sigma^2 \left[\alpha\beta + r(\beta^2 - 4\sigma^2)\right] \tau^5 + o(\tau^5)$$

as $\tau \to 0^+$. Now convergence is uniform w. r. to r on compact subintervals $[r_1, r_2] \subset [0, \infty)$.

4.3 Improved higher order approximation formula

In this section we present the main result of the paper [47] by Stehlíková and Ševčovič. It follows from (3) that the term $\ln P^{ap}(\tau, r) - c_5(r)\tau^5$ is the higher order accurate approximation of $\ln P^{ex}$ when compared to the original approximation $\ln P^{ap}(\tau, r)$ from [17]. Furthemore, we show, that it is even possible to compute $O(\tau^6)$ term and to obtain a new approximation $\ln P^{ap2}(\tau, r)$ such that the difference $\ln P^{ap2}(\tau, r) - \ln P^{ex}(\tau, r)$ is $o(\tau^6)$ for small values of $\tau > 0$.

Let P^{ex} be the exact bond price in the model (4.1). Let us define an improved approximation P^{ap2} by the formula

$$\ln P^{ap2}(\tau, r) = \ln P^{ap}(\tau, r) - c_5(r)\tau^5 - c_6(r)\tau^6 \tag{4.16}$$

[2] This is referred to as the relative mispricing in [17]

where $\ln P^{ap}$ is given by (4.3), $c_5(\tau)$ is given by (4.9) in Theorem 1 and

$$c_6(r) = \frac{1}{6}\left(\frac{1}{2}\sigma^2 r^{2\gamma} c_5''(r) + (\alpha + \beta r)c_5'(r) - k_5(r)\right)$$

where c_5' and c_5'' stand for the first and second derivative of $c_5(r)$ w. r. to r and k_5 is defined in (4.14).

Theorem 4. *[47, Theorem 4] The difference between the higher order approximation* $\ln P^{ap2}$ *given by (4.16) and the exact solution* $\ln P^{ex}$ *satisfies* $\ln P^{ap2}(\tau, r) - \ln P^{ex}(\tau, r) = o(\tau^6)$ *as* $\tau \to 0^+$. *Convergence is uniform w. r. to* r *on compact subintervals* $[r_1, r_2] \subset (0, \infty)$.

Proof: We have to prove that $g(\tau, r) = c_5(r)\tau^5 + c_6(r)\tau^6 + o(\tau^6)$ where c_5 and c_6 are given above. We already know the form of the coefficient $c_5 = c_5(r)$. Consider the following Taylor series expansions:

$$g(\tau, r) = \sum_{i=5}^{\infty} c_i(r)\tau^i, \quad h(\tau, r) = \sum_{i=4}^{\infty} k_i(r)\tau^i, \quad f(\tau, r) = \sum_{i=1}^{\infty} l_i(r)\tau^i.$$

The absolute term l_0 is zero because $f^{ex}(0, r) = \ln P^{ex}(0, r) = \ln 1 = 0$ for all $r > 0$. Substituting power series into equation (4.15) and comparing coefficients of the order τ^5 enables us to derive the identity:

$$-6c_6(r) + \frac{1}{2}\sigma^2 r^{2\gamma} c_5''(r) + (\alpha + \beta r)c_5'(r) - k_5(r) = 0$$

and hence

$$c_6(r) = \frac{1}{6}\left(\frac{1}{2}\sigma^2 r^{2\gamma} c_5''(r) + (\alpha + \beta r)c_5'(r) - k_5(r)\right).$$

The term $k_5(r)$ given by (4.14) is obtained by computing the expansion of h. $\diamondsuit$

The order of relative error of bond prices and order of error of interest rates for the new higher order approximation can be derived similarly as in Corollary 1.

Remark 3. *It is not obvious how to obtain the next higher order terms of expansion because the equations contain unknown coefficients* $l_i(r)$, $i \geq 1$, *of the logarithm of the exact solution which is not known explicitly.*

Remark 4. *In the case of the CIR model we have*

$$c_5^{CIR}(r) = -\frac{\sigma^2}{120}\left(\alpha\beta + r(\beta^2 - 4\sigma^2)\right), \; k_5^{CIR}(r) = \frac{\beta\sigma^2}{40}\left(\alpha\beta + (\beta^2 - 10\sigma^2)r\right)$$

and so

$$c_6^{CIR}(r) = \frac{\sigma^2}{360}\left(-2\alpha\beta^2 + 17\beta\sigma^2 r - 2\beta^3 r + 2\alpha\sigma^2\right).$$

Hence

$$\ln P^{ap2}_{CIR} = \ln P^{ap}_{CIR} + \frac{\sigma^2}{120}\left(\alpha\beta + r(\beta^2 - 4\sigma^2)\right)\tau^5 - \frac{\sigma^2}{360}\left(-2\alpha\beta^2 + 17\beta\sigma^2 r - 2\beta^3 r + 2\alpha\sigma^2\right)\tau^6$$

The theorem yields $\ln P^{ap2}_{CIR}(\tau, r) - \ln P^{ex}_{CIR}(\tau, r) = o(\tau^6)$. *By computing the expansions of both exact and this approximative solutions we finally obtain*

$$\ln P^{ap2}_{CIR}(\tau, r) = \ln P^{ex}_{CIR}(\tau, r) - \frac{\sigma^2}{5040}\Big(11\alpha\beta^3 + 11\beta^4 r - 34\alpha\beta\sigma^2 - 180\beta^2 r\sigma^2 + 34 r\sigma^4\Big)\tau^7 + o(\tau^7) \ \ as \ \ \tau \to 0^+.$$

Convergence is uniform w. r. to r *on compact subintervals* $[r_1, r_2] \subset [0, \infty)$.

4.4 Comparison of approximations to the exact solution for the CIR model

In this section we present a comparison of the original and improved approximations in the case of the CIR model where the exact solution is known. We use the parameter values from [17] and [47], i.e. $\alpha = 0.00315$, $\beta = -0.0555$ and $\sigma = 0.0894$.

In Table 4.1 we show L_∞ and L_2-norms[3] with respect to r of the difference $\ln P^{ap} - \ln P^{ex}$ and $\ln P^{ap2} - \ln P^{ex}$ where we considered $r \in [0, 0.15]$. We also compute the experimental order of convergence (EOC) in these norms. Recall that the experimental order of convergence gives an approximation of the exponent α of expected power law estimate for the error $\|\ln P^{ap}(\tau, .) - \ln P^{ex}(\tau, .)\| = O(\tau^\alpha)$ as $\tau \to 0^+$. The EOC_i is given by a ratio

$$EOC_i = \frac{\ln(err_i/err_{i+1})}{\ln(\tau_i/\tau_{i+1})} \quad \text{where} \quad err_i = \|\ln P^{ap}(\tau_i, .) - \ln P^{ex}(\tau_i, .)\|_p \, .$$

In Table 4.2 and Figure 4.1 we show the L_2-error of the difference between the original and improved approximations for larger values of τ. It turned out that the higher order approximation P^{ap2} gives about twice better approximation of bond prices in the long time horizon up to 10 years.

4.5 Properties of the approximate term structures

We consider only the case $\gamma > 0$, since for $\gamma = 0$ (the Vasicek model) the interest rates R^{ap} coincide with the exact ones and their properties are well known. The following

[3] L_p and L_∞ norms of a function f defined on a grid with step h are given by $\|f\|_p = \left(h \sum |f(x_i)|^p\right)^{1/p}$ and $\|f\|_\infty = \max |f(x_i)|$.

Table 4.1: The L_∞ and L_2-errors for the original $\ln P^{ap}_{CIR}$ and improved $\ln P^{ap2}_{CIR}$ approximations.

τ	$\|\ln P^{ap} - \ln P^{ex}\|_\infty$	EOC	$\|\ln P^{ap2} - \ln P^{ex}\|_\infty$	EOC
1	2.774×10^{-7}	4.930	4.682×10^{-10}	7.039
0.75	6.717×10^{-8}	4.951	6.181×10^{-11}	7.029
0.5	9.023×10^{-9}	4.972	3.576×10^{-12}	7.004
0.25	2.876×10^{-10}	–	2.786×10^{-14}	–

τ	$\|\ln P^{ap} - \ln P^{ex}\|_2$	EOC	$\|\ln P^{ap2} - \ln P^{ex}\|_2$	EOC
1	6.345×10^{-8}	4.933	9.828×10^{-11}	7.042
0.75	1.535×10^{-8}	4.953	1.296×10^{-11}	7.031
0.5	2.061×10^{-9}	4.973	7.492×10^{-13}	7.012
0.25	6.563×10^{-11}	–	5.805×10^{-15}	–

Table 4.2: The L_2-error with respect to r for large values of τ.

τ	1	2	3	4	5
$\|\ln P^{ap} - \ln P^{ex}\|_2$	6.345×10^{-8}	1.877×10^{-6}	1.314×10^{-5}	5.093×10^{-5}	1.427×10^{-4}
$\|\ln P^{ap2} - \ln P^{ex}\|_2$	9.828×10^{-11}	1.314×10^{-8}	2.329×10^{-7}	1.799×10^{-6}	8.798×10^{-6}

τ	6	7	8	9	10
$\|\ln P^{ap} - \ln P^{ex}\|_2$	3.255×10^{-4}	6.441×10^{-4}	1.148×10^{-3}	1.890×10^{-3}	2.921×10^{-3}
$\|\ln P^{ap2} - \ln P^{ex}\|_2$	3.217×10^{-5}	9.618×10^{-5}	2.479×10^{-4}	5.705×10^{-4}	1.200×10^{-3}

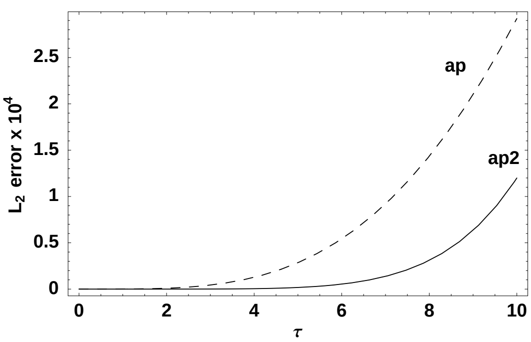

Figure 4.1: The error $\|\ln P^{ap}(\tau,.) - \ln P^{ex}(\tau,.)\|_2$ for the original approximation P^{ap} (dashed line) and the new approximation P^{ap2} (solid line). The horizontal axis is time τ to maturity.

theorem confirms a consistency of approximations. In particular, we are interested in limits of term structures as τ approaches zero. It is natural to require that they converge to the short rate r.

Theorem 5. *The limit of the interest rate $R^{ap}(\tau,r)$ and $R^{ap2}(\tau,r)$ equals r as $\tau \to 0^+$, i.e. the term structures are continuous in $\tau = 0$.*

Proof:

1. We write the interest rate R^{ap} in the following form:

$$\begin{aligned} R^{ap}(\tau,r) &= -\frac{1}{\tau}\ln P^{ap}(\tau,r) = r\frac{B}{\tau} - \frac{\alpha}{\beta}\frac{\tau - B}{\tau} - (r^{2\gamma} + q\tau)\frac{\sigma^2}{4\beta}\frac{B^2 + \frac{2}{\beta}(\tau - B)}{\tau} \\ & \quad + q\frac{\sigma^2}{8\beta^2}\frac{B^2(2\beta\tau - 1) - 2B\left(2\tau - \frac{3}{\beta}\right) + 2\tau^2 - \frac{6\tau}{\beta}}{\tau}. \end{aligned} \tag{4.17}$$

   Since

$$\lim_{\tau\to 0+} B(\tau) = \lim_{\tau\to 0+}\frac{e^{\beta\tau}-1}{\beta} = 0, \ \lim_{\tau\to 0+}\frac{B(\tau)}{\tau} = \lim_{\tau\to 0+}\frac{e^{\beta\tau}-1}{\beta\tau} = 1,$$

   we obtain

$$\begin{aligned} &\lim_{\tau\to 0+}\frac{\tau - B}{\tau} = 0, \\ &\lim_{\tau\to 0+}\frac{B^2 + \frac{2}{\beta}(\tau - B)}{\tau} = 0, \\ &\lim_{\tau\to 0+}\frac{B^2(2\beta\tau - 1) - 2B\left(2\tau - \frac{3}{\beta}\right) + 2\tau^2 - \frac{6\tau}{\beta}}{\tau} = 0, \end{aligned}$$

   from which the limit $\lim_{\tau\to 0^+} R^{ap}(\tau,r) = r$ in the theorem follows.

2. Since $\ln P^{ap2} = \ln P^{ap} + o(\tau^4)$, for interest rates we have

$$R^{ap2} = -\frac{\ln P^{ap} + o(\tau^4)}{\tau} = R^{ap} + o(\tau^3).$$

   As $\lim_{\tau\to 0+} R^{ap}(\tau,r) = r$, we also have $\lim_{\tau\to 0+} R^{ap2}(\tau,r) = r$ for any $r > 0$.

$\diamondsuit$

Infinite limits, derived in the next theorem, are one of the reasons, why the approximations, compared to exact CIR values, are not suitable for larger times to maturity. The other reason is that the expansions are done for $\tau \to 0^+$.

Theorem 6.

1. *The limit of interest rate R^{ap} for $\tau \to \infty$ is equal to minus infinity if $q < 0$, plus infinity if $q > 0$ and it has a finite limit $-\frac{\alpha}{\beta} + r^{2\gamma}\frac{\sigma^2}{2\beta^2}$ if $q(r) = 0$, where $q(r)$ is given by (4.4). A finite limit occurs for at most two positive values of r.*

2. *The condition $q = 0$ is necessary also for a finite limit of R^{ap2} as $\tau \to \infty$.*

Proof: Firstly we note that $B(\tau)/\tau \to 0$ as $\tau \to \infty$.

1. We use the expression (4.17) for interest rates. Then, for $\tau \to \infty$, we have a finite limit of the following part of the expression:

$$r\frac{B}{\tau} - \frac{\alpha}{\beta}\frac{\tau - B}{\tau} + q(r)\frac{\sigma^2}{8\beta^2}\frac{B^2(2\beta\tau - 1) - 2B\left(2\tau - \frac{3}{\beta} + 2\tau^2 - \frac{6\tau}{\beta}\right)}{\tau} \to -\frac{\alpha}{\beta}.$$

   The limit of the remaining part is

$$\lim_{\tau\to\infty} - \left((r^{2\gamma} + q(r)\tau)\frac{\sigma^2}{4\beta}\frac{B^2 + \frac{2}{\beta}(\tau - B)}{\tau}\right) = -r^{2\gamma}\frac{\sigma^2}{2\beta^2} - q\frac{3\sigma^2}{4\beta^3} - \frac{\sigma^2}{2\beta^2}\lim_{\tau\to\infty} q(r)\tau,$$

   from which the first statement of the theorem follows. What remains to be shown is that the equality $q = 0$ holds only at most for two positive value of r.

   This condition can be written as $\gamma(2\gamma - 1)\sigma^2 r^{2(2\gamma-1)} + 2\gamma r^{2\gamma-1}(\alpha + \beta r) = 0$ or, equivalently,

$$(2\gamma - 1)\sigma^2 r^{2\gamma-1} + 2(\alpha + \beta r) = 0, \tag{4.18}$$

   where we used the fact that γ is not zero and that $r > 0$. Denoting the left hand side of (4.18) by $f(r)$, we have $f'(r) = (2\gamma - 1)^2\sigma^2 r^{2\gamma-2} + 2\beta$. This derivative is equal to zero only for one positive r if $\gamma \neq 1/2$ and $f'(r) = 2\beta$ if $\gamma = 1/2$. Then, according to Rolle's theorem, there are at most two points in which $f(r) = 0$.

2. It follows from the expression $R^{ap2} = R^{ap} + c_5\tau^5 + c_6\tau^6$, since R^{ap} is not equal to any of the terms $-c_5\tau^5$, $-c_6\tau^6$, $-c_5\tau^5 - c_6\tau^6$ and hence it does not identically vanish.

$\diamondsuit$

Chapter 5

Calibration of one-factor models

In this chapter we propose a method for calibrating one-factor models using approximate analytical solutions from the previous chapter. We are interested in estimation of the parameter γ, i.e. the dependence of the volatility of the driving process of the short rate level. We suppose that

- under the real measure the short rate process is given by

$$dr = (a + br)dt + \sigma r^{\gamma} dw, \tag{5.1}$$

- under the risk-neutral measure the short rate process is given by

$$dr = (\alpha + \beta r)dt + \sigma r^{\gamma} dw, \tag{5.2}$$

i.e. the process has a linear drift under both real as well as risk neutral measures. Note that the volatilities are the same and the difference between drifts determines the market price of risk $\lambda(r) = \frac{(a-\alpha)+(b-\beta)r}{\sigma r^{\gamma}}$. Assumption (5.2) enables us to use approximate analytical solutions for interest rates studied in previous chapter and hence to consider deviations from interest rates on the real market. Assumption (5.1) enables us to use the Gaussian estimation methodology due to Nowman [32] for estimating parameters from time series of the short rate. The Gaussian methodology is based on a suitable approximation the likelihood function for $\gamma > 0$.

5.1 Nowman's Gaussian estimates

5.1.1 Discrete approximation of the model

In this section we show how the discrete approximation of the model is derived. The main idea is to approximate the volatility by a piecewise constant function, which stays constant between two observations (see [32], [23]).

The equation for the short rate under the real probability measure

$$dr_s = (\alpha + \beta r_s)ds + \sigma r_s^\gamma dw_s$$

is multiplied by $e^{-\beta s}$, which yields

$$e^{-\beta s} dr_s - \beta e^{-\beta s} r_s ds = \alpha e^{-\beta s} ds + \sigma e^{-\beta s} r_s^\gamma dw_s$$

$$\frac{d}{ds}\left(e^{\beta s} r_s\right) = \alpha e^{-\beta s} ds + \sigma e^{-\beta s} r_s^\gamma dw_s,$$

from which by integration over the time interval $[t-1, t]$ to time t we obtain

$$e^{-\beta t} r_t - e^{-\beta(t-1)} r_{t-1} = \frac{\alpha}{\beta}\left(e^{-\beta(t-1)} - e^{-\beta t}\right) + \int_{t-1}^{t} \sigma r_s^\gamma e^{-\beta s} dw_s.$$

Using the approximation, according to which the volatility is constant on interval $[t-1, t)$ and equals the values at the beginning of the interval, we get

$$\int_{t-1}^{t} \sigma r_s^\gamma e^{-\beta s} dw_s = \sigma r_{t-1}^\gamma \int_{t-1}^{t} e^{-\beta s} dw_s,$$

and hence

$$e^{-\beta t} r_t - e^{-\beta(t-1)} r_{t-1} = \frac{\alpha}{\beta}\left(e^{-\beta(t-1)} - e^{-\beta t}\right) + \sigma r_{t-1}^\gamma \int_{t-1}^{t} e^{-\beta s} dw_s.$$

Multiplying this equation by the term $e^{\beta t}$ and denoting

$$\varepsilon_t = \sigma r_{t-1}^\gamma e^{\beta t} \int_{t-1}^{t} e^{-\beta s} dw_s$$

we obtain a discrete short rate model

$$r_t = e^{\beta} r_{t-1} + \frac{\alpha}{\beta}\left(e^{\beta} - 1\right) + \varepsilon_t \quad \text{for } t = 2, \ldots, N. \tag{5.3}$$

The onditional distribution of ε_t for a given value of r_{t-1} follows from properties of Itō's integral (c.f. [37]): ε_t are normally distributed, uncorrelated for $t = 1, 2, \ldots$, with a zero expected value and the variance ν_t^2 satisfying

$$\begin{aligned} \nu_t^2 &:= Var(\varepsilon_t) = \sigma^2 r_{t-1}^{2\gamma} e^{2\beta t}\, Var\left(\int_{t-1}^{t} e^{-\beta s} dw_s\right) \\ &= \sigma^2 r_{t-1}^{2\gamma} \int_{t-1}^{t} e^{-2\beta s} ds = \sigma^2 r_{t-1}^{2\gamma} \frac{e^{2\beta} - 1}{2\beta}, \end{aligned}$$

where we used Itō's isometry (Corollary 3.1.7, [37]).

The likelihood function L for this model is equal (up to an additive constant) to (see [32], [23])

$$\log L = -\frac{1}{2}\sum_{t=2}^{N}\left(\log \nu_t^2 + \frac{\varepsilon_t^2}{\nu_t^2}\right), \tag{5.4}$$

where

$$\nu_t^2 = \frac{\sigma^2}{2\beta}\left(e^{\beta}-1\right) r_{t-1}^{2\gamma},\ \varepsilon_t = r_t - \frac{\alpha}{\beta}\left(e^{\beta}-1\right) - e^{\beta} r_{t-1}. \tag{5.5}$$

Estimates of the parameters are arguments of the maximum of the function $\log L$.

These results are obtained in the case when the interval $[t-1,t]$ between two consecutive values of r_t is taken as a unit of time. In a case of another time scale when the length of the interval is Δt, in the same way as before we derive the model

$$r_k = e^{\beta\Delta t} r_{k-1} + \frac{\alpha}{\beta}\left(e^{\beta\Delta t}-1\right) + \varepsilon_k \ \ (k=2,\ldots,N), \tag{5.6}$$

where k is a number of observations (to simplify the notation, we index the observations by their number, instead of time), with ε_k normally distributed, uncorrelated, with a zero expected value and the variance $\sigma^2 r_{k-1}^{2\gamma}\frac{e^{2\beta\Delta t}-1}{2\beta}$. This can be written as

$$r_k = e^{\tilde{\beta}} r_{k-1} + \frac{\tilde{\alpha}}{\tilde{\beta}}\left(e^{\tilde{\beta}}-1\right) + \tilde{\varepsilon}_k \ \ (k=2,\ldots,N), \tag{5.7}$$

where $\tilde{\varepsilon}_k$ are normally distributed, uncorrelated, with a zero expected value and the variance $\tilde{\sigma}^2 r_{k-1}^{2\gamma}\frac{e^{2\tilde{\beta}}-1}{2\tilde{\beta}}$, where

$$\tilde{\alpha} = \alpha\Delta t,\ \tilde{\beta} = \beta\Delta t, \tilde{\sigma^2} = \sigma^2\Delta t. \tag{5.8}$$

When studying the existence of maximum of the likelihood function, we can study model in the form (5.7), which is equivalent to (5.3), and when estimating the parameters α, β, σ^2, we divide estimates of $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\sigma}^2$ by Δt (resp., when estimating σ, we divide $\tilde{\sigma}^2$ by $\sqrt{\Delta t}$).

5.1.2 Examples of calibration

We present examples of calibration results using the Gaussian methodology. We use Bribor overnight daily data from 2007. Their evolution over the year is shown in the Figure 5.1 and their basic descriptive statistics are shown in the Figure 5.2.

Figures 5.3 and 5.4 show drift and volatility functions for several values of γ, estimated from Bribor overnight interest rates in 2007. Table 5.1 presents numerical values of the estimates.

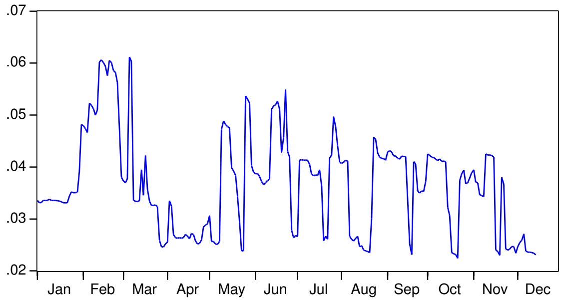

Figure 5.1: Bribor overnight (short rate) process, 2007.

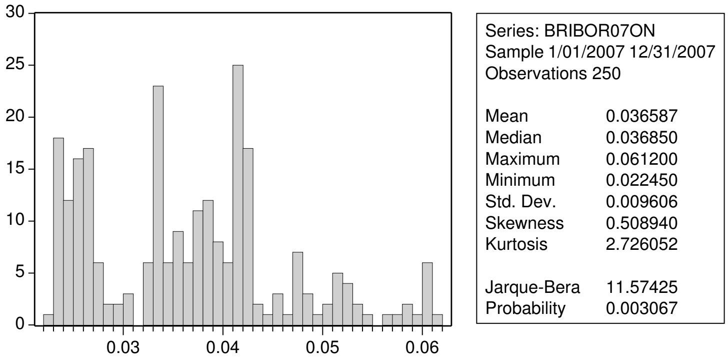

Figure 5.2: Descriptive statistics of Bribor overnight, 2007.

Table 5.1: Gaussian estimates for Bribor, 2007.

γ	α	β	σ
0	1.5261	-41.9624	0.0888
0.5	1.6370	-44.9889	0.4917
1	1.7762	-49.0323	2.8446
1.5	1.9465	-54.2752	17.0548

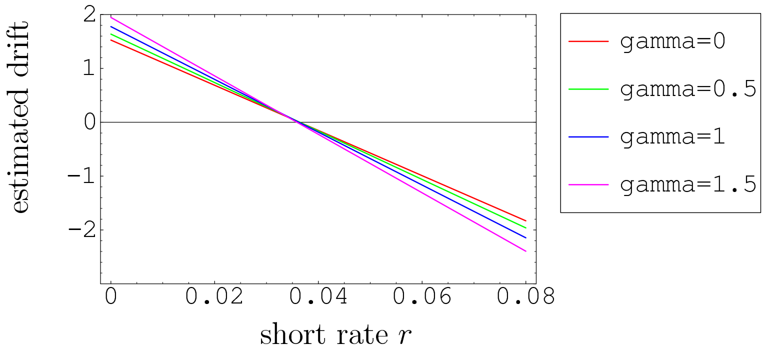

Figure 5.3: Estimated drift for several values of γ (Bribor, 2007).

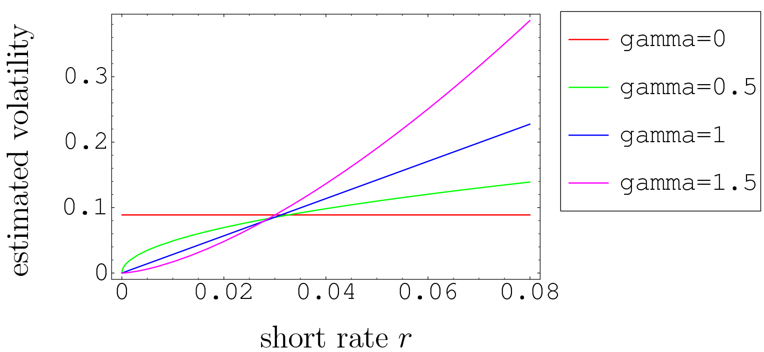

Figure 5.4: Estimated volatility for several values of γ (Bribor, 2007).

5.1.3 The condition for existence of the log-likelihood function maximum

Using the following transformation:

$$a = \frac{\alpha}{\beta}(e^{\beta} - 1),\ b = e^{\beta},\ s^2 = \sigma^2 \frac{e^{2\beta} - 1}{2\beta} \tag{5.9}$$

model (5.3) can be transformed into the form $r_t = a + br_{t-1} + \varepsilon_t$, where ε_t are uncorrelated and conditional distribution of $\varepsilon_t | r_{t-1}$ is $N(0, s^2 r_{t-1}^{2\gamma})$. In general, the estimate of b need not be positive. But if it is not positive, we cannot do inverse transformation and obtain original parameters. Now we investigate when the maximum of the likelihood function does not exist.

We return to original parameters α, β, σ^2 of the model. Recall that the likelihood function L for this model is equal (up to an additive constant) to

$$\log L = -\frac{1}{2} \sum_{t=2}^{N} \left(\log \nu_t^2 + \frac{\varepsilon_t^2}{\nu_t^2} \right), \tag{5.10}$$

where

$$\nu_t^2 = \frac{\sigma^2}{2\beta} \left(e^{\beta} - 1 \right) r_{t-1}^{2\gamma},\ \varepsilon_t = r_t - \frac{\alpha}{\beta} \left(e^{\beta} - 1 \right) - e^{\beta} r_{t-1}. \tag{5.11}$$

We fix γ and look for the maximum of $\log L$ with respect to parameters α, β and σ^2.

Since

$$\frac{\partial \log L}{\partial \alpha} = \sum_{t=2}^{N} \frac{\partial \log L}{\partial \varepsilon_t} \frac{\partial \varepsilon_t}{\partial \alpha} = \frac{e^{\beta} - 1}{\beta} \sum_{t=2}^{N} \frac{\varepsilon_t}{\nu_t^2},$$

from necessary first order condition $\frac{\partial \log L}{\partial \alpha}\Big|_{(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2)} = 0$ we obtain

$$\sum_{t=2}^{N} \frac{\hat{\varepsilon}_t}{\hat{\nu}_t^2} = 0,$$

where $\hat{\varepsilon}_t, \hat{\nu}_t^2$ are values ε_t, ν_t^2 evaluated in the points $\left(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2 \right)$. After substituting $\hat{\nu}_t^2$ we obtain the equality

$$\sum_{t=2}^{N} \hat{\varepsilon}_t r_{t-1}^{2\gamma} = 0.$$

inserting $\hat{\varepsilon}_t$ into the above equality, we are able to obtain the estimate of the parameter α in the form:

$$\hat{\alpha} = \frac{\hat{\beta}}{e^{\hat{\beta}} - 1} \frac{\sum_{t=2}^{N} r_t r_{t-1}^{-2\gamma} - e^{\hat{\beta}} \sum_{t=2}^{N} r_{t-1}^{1-2\gamma}}{\sum_{t=2}^{N} r_{t-1}^{-2\gamma}}. \tag{5.12}$$

Similarly, since

$$\frac{\partial \log L}{\partial \sigma^2} = \sum_{t=2}^{N} \frac{\partial \log L}{\partial \nu_t^2}\frac{\partial \nu_t^2}{\partial \sigma^2} = \sum_{t=2}^{N}\left(-\frac{1}{2\nu_y^2} + \frac{\varepsilon_t^2}{2(\nu_t^2)^2}\right)\frac{e^{2\beta}-1}{2\beta}r_{t-1}^{2\gamma}$$

and $\frac{\partial \log L}{\partial \sigma^2}\big|_{(\hat{\alpha},\hat{\beta},\hat{\sigma}^2)} = 0$, we get

$$\sum_{t=2}^{N}\left(\frac{1}{\hat{\nu}_t^2} - \frac{\hat{\varepsilon}_t^2}{(\hat{\nu}^2)^2}\right) r_{t-1}^{2\gamma} = 0.$$

From this, after substituting $\hat{\varepsilon}_t$ and $\hat{\nu}_t^2$, we have

$$\hat{\sigma}^2\frac{e^{2\hat{\beta}}-1}{2\hat{\beta}}(N-1) - \sum_{t=2}^{N}\hat{\varepsilon}_t^2 r_{t-1}^{2\gamma} = 0.$$

Hence, we derived a formula for $\hat{\sigma}^2$:

$$\hat{\sigma}^2 = \frac{1}{N-1}\frac{2\hat{\beta}}{e^{2\hat{\beta}}-1}\sum_{t=2}^{N}\hat{\varepsilon}_t^2 r_{t-1}^{2\gamma}. \tag{5.13}$$

Let us define the following functions, based on formulae (5.12) and (5.13):

$$\alpha(\beta) = \frac{\beta}{e^{\beta}-1}\frac{\sum_{t=2}^{N} r_t r_{t-1}^{-2\gamma} - e^{\beta}\sum_{t=2}^{N} r_{t-1}^{1-2\gamma}}{\sum_{t=2}^{N} r_{t-1}^{-2\gamma}}, \tag{5.14}$$

$$\sigma^2(\beta) = \frac{1}{N-1}\frac{2\beta}{e^{2\beta}-1}\sum_{t=2}^{N}\varepsilon_t^2 r_{t-1}^{2\gamma} = 0, \tag{5.15}$$

where $\varepsilon_t = \varepsilon_t(\beta) = r_t - e^{\beta}r_{t-1} - \alpha(\beta)\frac{e^{\beta}-1}{\beta}$ and insert them for α and σ^2 in the definition of $\log L$. In this way, we obtain a function of one variable (since γ is given) $\log \tilde{L}(\beta)$. Now, we find its maximum and investigate the condition for its existence.

Substituting (5.15) into ν_t^2 we have

$$\nu_t^2 = \frac{1}{N-1}\left(\sum_{s=2}^{N}\varepsilon_s^2 r_{s-1}^{2\gamma}\right) r_{t-1}^{2\gamma}.$$

It follows that

$$\sum_{t=2}^{N}\log \nu_t^2 = (N-1)\log\frac{1}{N-1} + (N-1)\log\sum_{t=2}^{N}\varepsilon_t^2 r_{t-1}^{2\gamma} + 2\gamma\sum_{t=2}^{N}\log r_{t-1}$$

and

$$\sum_{t=2}^{N}\frac{\varepsilon_t^2}{\nu_t^2} = \sum_{t=2}^{N}\frac{\varepsilon_t^2}{\frac{1}{N-1}\left(\sum_{s=2}^{N}\varepsilon_s^2 r_{s-1}^{-2\gamma}\right) r_{t-1}^{2\gamma}} = N-1.$$

It means that $\log \tilde{L}(\beta)$ can be written as

$$-\frac{1}{2}\left((N-1)\log\frac{1}{N-1}+(N-1)\log\sum_{t=2}^{N}\varepsilon_t^2 r_{t-1}^{-2\gamma}+2\gamma\sum_{t=2}^{N}\log r_{t-1}+(N-1)\right),$$

where the only term depending on β is $-\frac{1}{2}(N-1)\log\sum_{t=2}^{N}\varepsilon_t^2 r_{t-1}^{-2\gamma}$. Hence it suffices to minimize the sum $\sum_{t=2}^{N}\varepsilon_t^2 r_{t-1}^{-2\gamma}$.

Inserting (5.14) into ε_t, we have $\varepsilon_t = A_t + e^{\beta}B_t$, where

$$A_t = r_t - \frac{\sum_{s=2}^{N} r_s r_{s-1}^{-2\gamma}}{\sum_{s=2}^{N} r_{s-1}^{-2\gamma}},\ B_t = \frac{\sum_{s=2}^{N} r_{s-1}^{1-2\gamma}}{\sum_{s=2}^{N} r_{s-1}^{-2\gamma}} - r_{t-1} \tag{5.16}$$

are independent of β. Then

$$\sum_{t=2}^{N}\varepsilon_t^2 r_{t-1}^{-2\gamma} = \left(\sum_{t=2}^{N} A_t^2 r_{t-1}^{-2\gamma}\right) + \left(\sum_{t=2}^{N} 2A_tB_t r_{t-1}^{-2\gamma}\right)e^{\beta} + \left(\sum_{t=2}^{N} B_t^2 r_{t-1}^{-2\gamma}\right)e^{2\beta}.$$

Minimization of $\log L$ is therefore equivalent to minimization of the function $f(\beta) = k_1e^{\beta} + k_2e^{2\beta}$, where

$$k_1 = \sum_{t=2}^{N} 2A_tB_t r_{t-1}^{-2\gamma},\ k_2 = \sum_{t=2}^{N} B_t^2 r_{t-1}^{-2\gamma}. \tag{5.17}$$

We have $k_2 \geq 0$ and if $r_1,\ldots,r_{N-1}$ are not all identically equal, we have the strict inequality[1] $k_2 > 0$. Further, we consider only this case. The function $f(\beta)$ then depends on the sign of k_1:

- If $k_1 \geq 0$, then $f(\beta) > 0$ for all β and $\lim_{\beta\to-\infty} f(\beta) = 0$. It means that the function $f(\beta)$ does not have a minimum. It converges to its infimum as $\beta \to -\infty$.
- If $k_1 < 0$, then f decreases for $\beta < \log\left(-\frac{k_1}{2k_2}\right)$ and increases for $\beta > \log\left(-\frac{k_1}{2k_2}\right)$. Thus, at the point $\beta = \log\left(-\frac{k_1}{2k_2}\right)$ the function f has its global minimum.

From (5.17) we get

$$k_1 = 2\,\frac{\left(\sum_{t=2}^{N} r_t r_{t-1}^{-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{1-2\gamma}\right) - \left(\sum_{t=2}^{N} r_t r_{t-1}^{1-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right)}{\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right)},$$

[1] $k_2 = 0$ if and only if all B_t $(t = 2,\ldots,N)$ are zero, i.e.

$$\frac{\sum_{s=2}^{N} r_{s-1}^{1-2\gamma}}{\sum_{s=2}^{N} r_{s-1}^{-2\gamma}} = r_{t-1}\ \text{ for } t = 2,\ldots,N.$$

The term on the left hand side in constant with respect to t, and hence r_{t-1} for $t = 2,\ldots,N$ (i.e. r_t for $t = 1,\ldots N-1$) are identically equal.

$$k_2 = \frac{\left(\sum_{t=2}^{N} r_{t-1}^{2-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right) - \left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right)^2}{\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right)}.$$

Hence the condition $k_1 < 0$ for existence of the minimum of f is equivalent to the statement

$$\left(\sum_{t=2}^{N} r_t r_{t-1}^{-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{1-2\gamma}\right) - \left(\sum_{t=2}^{N} r_t r_{t-1}^{1-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right) < 0. \tag{5.18}$$

In summary, we have shown the following statements:

Theorem 7.

- *If the inequality (5.18) holds, then the estimate of parameter β is*

$$\hat{\beta} = \log \frac{\left(\sum_{t=2}^{N} r_t r_{t-1}^{-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{1-2\gamma}\right) - \left(\sum_{t=2}^{N} r_t r_{t-1}^{1-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right)}{\left(\sum_{t=2}^{N} r_{t-1}^{2-2\gamma}\right)\left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right) - \left(\sum_{t=2}^{N} r_{t-1}^{-2\gamma}\right)^2},$$

  and estimates of parameters α a σ^2 are given by (5.14) and (5.15), evaluated for $\beta = \hat{\beta}$.

- *If inequality (5.18) does not hold, then maximum of likelihood function does not exist. The likelihood function is increasing on the curve parameterized by β for $\beta \to -\infty$: $(\alpha, \beta, \sigma^2) = (\alpha(\beta), \beta, \sigma^2(\beta))$ where $\alpha(\beta)$ and $\sigma^2(\beta)$ are given by (5.14) and (5.15).*

5.1.4 A theoretical example of nonexistence of maximum of the log-likelihood function

As we have shown in the previous section, it can happen that the likelihood function L does not attain its maximum and it is increasing on a curve where $\beta \to -\infty$. Writing the drift in the form $\kappa(\theta - r)$, it corresponds to the limit $\kappa \to \infty$. Hence, we could expect this to happen, when there is an evidence for a very strong mean reversion. Using this idea, we construct an example, where maximum likelihood estimates of the Vasicek model do not exist.

Let a, b be positive constants. Define

$$r_t = a + b\frac{(-1)^t}{t}, \quad t = 1, 2, \ldots, N. \tag{5.19}$$

An example of a sequence of r_t, obtained in this way, is shown in Figure 5.5. We show that, for any a and b, maximum likelihood estimate of the Vasicek model does not exist, i.e. the condition (for $\gamma = 0$)

$$\left(\sum_{t=2}^{N} r_t\right)\left(\sum_{t=2}^{N} r_{t-1}\right) - (N-1)\left(\sum_{t=2}^{N} r_t r_{t-1}\right) < 0.$$

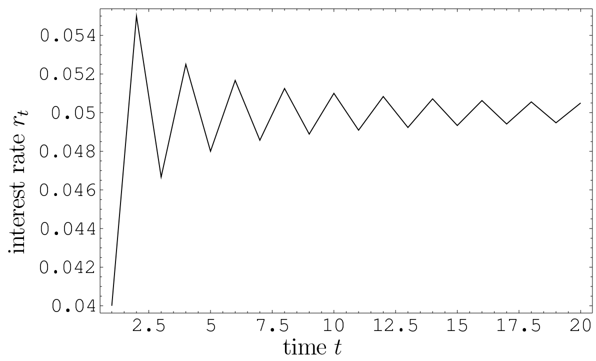

Figure 5.5: A sequence of r_t obtained from (5.19).

	j=1	j=2	j=3	j=4	j=5
i=2		+	-	+	-
i=3	+		+	-	+
i=4	-	+		+	-
i=5	+	-	+		+
i=6	-	+	-	+	

Figure 5.6: Graphical hint for the proof of the inequality (5.20).

is not satisfied. Substituting (5.19) into the left hand side of this inequality, yields

$$\left(\sum_{t=2}^{N} a + b\frac{(-1)^t}{t}\right)\left(\sum_{t=2}^{N} a + b\frac{(-1)^{t-1}}{t-1}\right) - (N-1)\sum_{t=2}^{N}\left(a + b\frac{(-1)^t}{t}\right)\left(a + b\frac{(-1)^{t-1}}{t-1}\right) =$$

$$= b^2\left(\sum_{t=2}^{N}\frac{(-1)^t}{t}\sum_{t=2}^{N}\frac{(-1)^{t-1}}{t-1} + \sum_{t=2}^{N}\frac{1}{t(t-1)}\right).$$

Now, we show that

$$\sum_{t=2}^{N}\frac{(-1)^t}{t}\sum_{t=2}^{N}\frac{(-1)^{t-1}}{t-1} + \sum_{t=2}^{N}\frac{1}{t(t-1)} > 0, \tag{5.20}$$

from which our claim about nonexistence a of maximum of the likelihood function follows.

The first term in the inequality (5.20), i.e. $\sum_{t=2}^{N}\frac{(-1)^t}{t}\sum_{t=2}^{N}\frac{(-1)^{t-1}}{t-1}$, is a sum of $\frac{(-1)^i}{i}\frac{(-1)^j}{j}$, where $i = 2, 3, \ldots, N$ a $j = 1, 2, \ldots, N-1$. Product $\frac{1}{ij}$ hence is hence included with positive sign if $i+j$ is even and with negative sign if $i+j$ is odd. Hence the left hand side of 5.20 equals

$$\begin{aligned}
&\sum_{t=2}^{N} \frac{(-1)^t}{t} \sum_{t=2}^{N} \frac{(-1)^{t-1}}{t-1} + \sum_{t=2}^{N} \frac{1}{t(t-1)} \\
&= \sum_{i=2}^{N} \sum_{j=1}^{N-1} \frac{(-1)^{i+j}}{ij} + \sum_{t=2}^{N} \frac{1}{t(t-1)} = \sum_{i=2}^{N} \sum_{j=1, j\neq i}^{N-1} \frac{(-1)^{i+j}}{ij} \qquad (5.21)
\end{aligned}$$

This summation procedure is illustrated in Figure 5.6. Diagonal elements are not included, as the sum (5.21) runs for $j \neq i$. Signs in the cells are the signs, by which the product $1/ij$ is included in 5.21.

Let us consider the following sums, each of which consists of the two terms $\frac{(-1)^{i+j}}{ij}$ from the sum 5.21. We distinguish

- sums below the diagonal:

$$\frac{1}{i(i-2)} - \frac{1}{(i+1)(i-2)} = \frac{1}{i(i+1)(i-2)} > 0$$

  (as $i \geq 3$ in this case),

- sums above the diagonal:

$$\frac{1}{i^2} - \frac{1}{i(i+1)} = \frac{1}{i^2(i+1)} > 0.$$

All the remaining terms are positive. Therefore the whole sum in (5.21) is positive.

Let us note that the inequality (5.20) for large N can be alternatively derived as follows. Since

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln 2,$$

we can compute the limit of the left hand side of (5.20):

$$\begin{aligned}
&\lim_{N\to\infty} \left(\sum_{t=2}^{N} \frac{(-1)^t}{t} \sum_{t=2}^{N} \frac{(-1)^{t-1}}{t-1} + \sum_{t=2}^{N} \frac{1}{t(t-1)} \right) = \\
&\lim_{N\to\infty} \left(\left[1 - \sum_{t=1}^{N-1} \frac{(-1)^{t+1}}{t} \right] \left[- \sum_{t=1}^{N-1} \frac{(-1)^{t+1}}{t} \right] + \sum_{t=2}^{N} \left[\frac{1}{t-1} - \frac{1}{t} \right] \right) = \\
&\lim_{N\to\infty} \left(\left[1 - \sum_{t=1}^{N-1} \frac{(-1)^{t+1}}{t} \right] \left[- \sum_{t=1}^{N-1} \frac{(-1)^{t+1}}{t} \right] + 1 - \frac{1}{N} \right) = \\
&(1 - \ln 2)(-\ln 2) + 1 \approx 0.7873 > 0.
\end{aligned}$$

This can be considered as an alternative proof of inequality (5.20) for sufficiently large N.

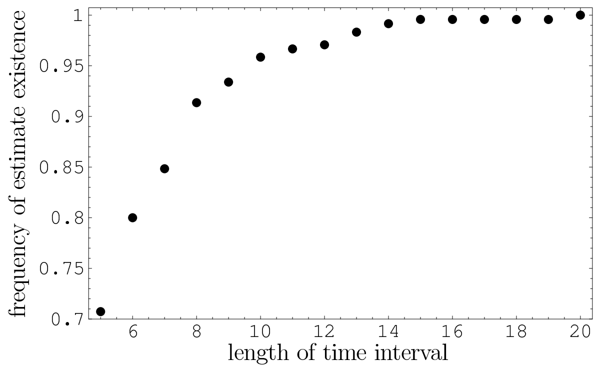

Figure 5.7: Frequency of time intervals, in which Gaussian estimates of the Vasicek model ($\gamma = 0$) exist (Bribor, 2007).

Table 5.2: Frequency of time intervals, in which Gaussian estimates exist (Bribor, 2007).

length of time interval	$\gamma = 0$	$\gamma = 0.5$	$\gamma = 1$
5	0.7073	0.7033	0.7073
10	0.9585	0.9461	0.9419
15	0.9958	0.9915	0.9915
20	1.0000	1.0000	0.9957

5.1.5 Existence and nonexistence of the log-likelihood function maximum for real market data

Now we consider real data and the question of existence of Gaussian estimates. We take time intervals of different lengths. If the length is k days, we take data from intervals $[1, k]$, $[2, k+1]$, etc. and we compute a frequency of the instances when Gaussian estimate exists. For the Vasicek model, the results are shown in Figure 5.7. Similar results are obtained also for another values of γ. They are summarized in Table 5.2.

It can be observed that the problem of nonexistence of the maximum arises especially for small number of observations. For example, in the Vasicek model, the maximum exists only in 70.73 percent of the cases when we use 5 observations (i.e. one week). When we use 10 observations, it is already 95.85 percent. For 20 observations (approximately one month), estimates exist in all the cases.

5.2 Comparison with whole term structures

Following the approach developed in [48] and [49], our aim is to minimize the differences between real data and interest rates predicted by a model. The objective function is

$$F = \sum_{i,j} w_{ij}(R(r_i, \tau_j) - R_{ij})^2,$$

where we index time by i and times to maturity of interest rates by j. $R(\tau_j, r_i)$ is an interest rate with maturity τ_j given by a model, when the short rate equals r_i, R_{ij} is a real observed interest rate. This function is minimized with respect to parameters of the model.

The weight functions w_{ij} provide the weight to the differences. We can prescribe different weights for older data and more recent data (by making w dependent on i), or for interest rates with different maturities (by making it dependent on j).

5.3 Proposed method of calibration and results

It is known that Gaussian estimates of parameters α and β are biased, while σ is estimated more precisely, see for example simulation study [52]. Moreover, the drift functions in real probability (which was used in Gaussian estimates) and in risk neutral measure (which is needed to compute interest rates) are not identical. Since the volatility does not change with this change of probability measure, one possible way of calibration (suggested e.g. in [12] for the Vasicek model) is to estimate the volatility by Gaussian methodology and then use the interest rates to estimate the risk neutral drift.

Following [48], [49] we consider the objective function for comparison of theoretical and real interest rates

$$\begin{aligned} F(\alpha, \beta) &= \sum_{i,j} w_{ij}(R(r_i, \tau_j) - R_{ij})^2 \\ &= \sum_{i,j} w_{ij} \left(-\frac{\ln P^{ap}(r_i, \tau_j)}{\tau_j} - R_{ij} \right)^2 = \sum_{i,j} \frac{w_{ij}}{\tau_j^2} \left(-\tau_j R_{ij} - \ln P^{ap}(r_i, \tau_j) \right)^2 \\ &= \sum_{i,j} \frac{w_{ij}}{\tau_j^2} \left(\tau_j R_{ij} + \ln P^{ap}(r_i, \tau_j) \right)^2 . \end{aligned} \tag{5.22}$$

Notice that the theoretical values $R(r_i, \tau_j)$ are calculated using approximate analytical solution P^{ap} discussed in chapter 3. Compared to numerical computation of the interest rates based on a solution to PDE (4.2), this approach has the advantage of having high accuracy obtained with a faster computation of P^{ap} in a closed form. Recall that

$$\begin{aligned} \ln P^{ap}(\tau, r) &= -rB + \frac{\alpha}{\beta}(\tau - B) + \left(r^{2\gamma} + q\tau\right) \frac{\sigma^2}{4\beta} \left[B^2 + \frac{2}{\beta}(\tau - B) \right] \\ &\quad -q\frac{\sigma^2}{8\beta^2} \left[B^2(2\beta\tau - 1) - 2B\left(2\tau - \frac{3}{\beta}\right) + 2\tau^2 - \frac{6\tau}{\beta} \right] \end{aligned}$$

where

$$q(r) = \gamma(2\gamma-1)\sigma^2 r^{2(2\gamma-1)} + 2\gamma r^{2\gamma-1}(\alpha+\beta r)$$

and $B(\tau) = (e^{\beta\tau}-1)/\beta$, see (4.3) and (4.4). Note that $q(r)$ and $\ln P^{ap}(\tau,r)$ are linear functions of α. They can be written as

$$\begin{aligned} q(r) &= q_1(r), \\ \ln P^{ap}(r,\tau) &= c_1 + \alpha c_2 \end{aligned}$$

where

$$\begin{aligned} c_1 &= c_1(r,\tau,\beta,\sigma,\gamma) = -rB + \frac{\sigma^2}{4\beta}\left(B^2 + \frac{2}{\beta}(\tau - B)\right)\left(r^{2\gamma} + q_2\tau\right) \\ &\quad -q_2\frac{\sigma^2}{8\beta^2}\left(B^2(2\beta\tau-1) - 2B\left(2\tau - \frac{3}{\beta}\right) + 2\tau^2 - \frac{6\tau}{\beta}\right), \\ c_2 &= c_2(r,\tau,\beta,\sigma,\gamma) = \frac{\tau - B}{\beta} + q_1\tau\frac{\sigma^2}{4\beta}\left(B^2 + \frac{2}{\beta}(\tau - B)\right) \\ &\quad -q_1\frac{\sigma^2}{8\beta^2}\left(B^2(2\beta\tau-1) - 2B\left(2\tau - \frac{3}{\beta}\right) + 2\tau^2 - \frac{6\tau}{\beta}\right) \end{aligned}$$

and

$$\begin{aligned} q_1 &= q_1(r,\gamma) = 2\gamma r^{2\gamma-1}, \\ q_2 &= q_2(r,\gamma,\sigma) = \gamma(2\gamma-1)\sigma^2 r^{3(3\gamma-1)} + 2\gamma r^{2\gamma}. \end{aligned}$$

Substituting them into (5.22), we write the function $F(\alpha,\beta)$ as a quadratic function of α:

$$F(\alpha,\beta) = \left[\sum_{i,j}\frac{w_{ij}}{\tau_j^2}c_2^2\right]\alpha^2 + \left[2\sum_{i,j}\frac{w_{ij}}{\tau_j^2}c_2(\tau_j R_{ij} + c_1)\right]\alpha + \left[\sum_{i,j}\frac{w_{ij}}{\tau_j^2}(\tau_j R_{ij} + c_1)^2\right]. \tag{5.23}$$

Hence, when the parameter β is prescribed, the corresponding optimal value of $\alpha = \alpha(\beta)$ is given by

$$\alpha(\beta) = -\frac{\sum_{i,j}\frac{w_{ij}}{\tau_j^2}c_2(\tau_j R_{ij} + c_1)}{\sum_{i,j}\frac{w_{ij}}{\tau_j^2}c_2^2}. \tag{5.24}$$

Substituting (5.24) into (5.23) we can write F as a function of the one variable β, i.e.

$$F(\beta) = -\frac{\left[\sum_{i,j}\frac{w_{ij}}{\tau_j^2}c_2(\tau_j R_{ij} + c_1)\right]^2}{\sum_{i,j}\frac{w_{ij}}{\tau_j^2}c_2^2} + \sum_{i,j}\frac{w_{ij}}{\tau_j^2}(\tau_j R_{ij} + c_1)^2. \tag{5.25}$$

To find the solution of our optimization problem, this function has to be minimized with respect to β. The optimal value of α is then given by (5.24).

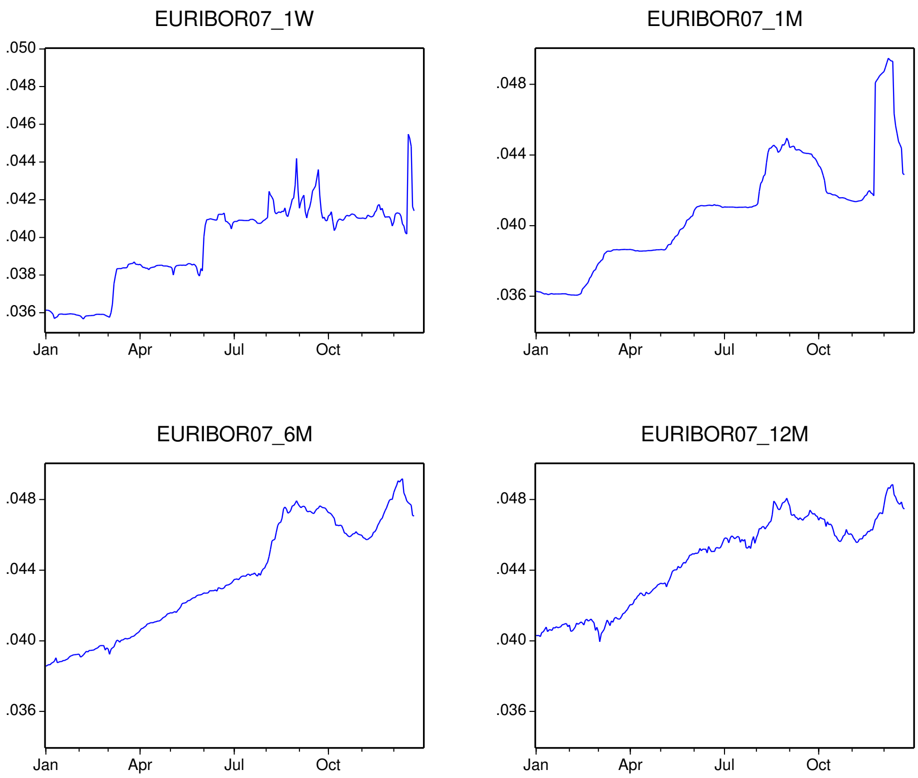

Figure 5.8: Euribor data for selected maturities, 2007.

We show results of the calibration for Euribor interest rates. We use daily data from the year 2007. As an approximation for the short rate we use 1 week interest rate and we use 12 maturities for term structures, from 1 month up to 12 month maturity. Plots of data for selected maturities are shows in Figure 5.8.

We consider two weight functions: $w_{ij} = \tau_j^2$ (i.e. we put higher weigh on longer maturities) and $w_{ij} = 1/\tau_j^2$ (i.e. we put higher weight on shorter maturities). In Figure 5.9, the function $F(\beta)$ is shown for $w_{ij} = \tau_j^2$ and $\gamma = 0$. It has a similar shape for other choices of w_{ij} and γ.

Tables 5.3 and 5.4 present the estimates of parameters for several values of γ. Figure 5.13 shows the attained optimal values of objective function. The optimal γ is significantly influenced by the choice of weights w. In case of $w_{ij} = \tau_j^2$, the Vasicek model is confirmed, the smallest value of the objective function is attained for $\gamma = 0$. However, when using $w_{ij} = 1/\tau_j^2$, the optimal value of γ is found to be close to 3.

Estimates of σ are obtained from the short rate by Gaussian methodology. Hence they are independent of weights. They have similar behavior as for the Bribor short rate - the estimated volatility functions intersect with each other for similar values of short rate, see Figure 5.10. Drifts estimated by both weights are shown in the Figure

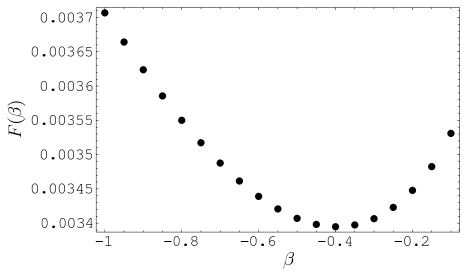

Figure 5.9: The plot of the function $F(\beta)$.

Table 5.3: Estimates for weights $w_{ij} = \tau_j^2$.

γ	α	β	σ
0	0.0162	-0.3781	0.0341
0.5	0.0164	-0.3959	0.1638
1	0.0182	-0.4552	0.7877
1.5	0.0231	-0.5918	3.7931
2	0.0326	-0.8455	18.2938
2.5	0.0493	-1.2789	88.3684
3	0.0771	-1.9888	427.526
3.5	0.1236	-3.1637	2071.53

Table 5.4: Estimates for weights $w_{ij} = 1/\tau_j^2$.

γ	α	β	σ
0	0.1608	-3.7277	0.0341
0.5	0.1605	-3.7306	0.1638
1	0.1613	-3.7600	0.7877
1.5	0.1634	-3.8194	3.7931
2	0.1671	-3.9180	18.2938
2.5	0.1732	-4.0759	88.3684
3	0.1833	-4.3316	427.526
3.5	0.2002	-4.7489	2071.53

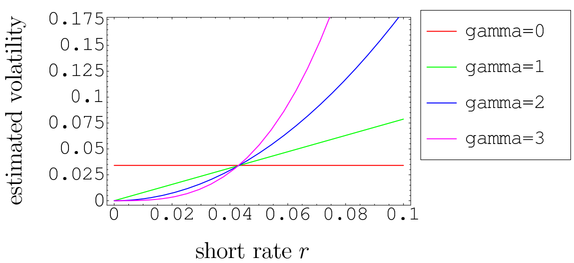

Figure 5.10: Estimated volatility σr^{γ} (Euribor, 2007).

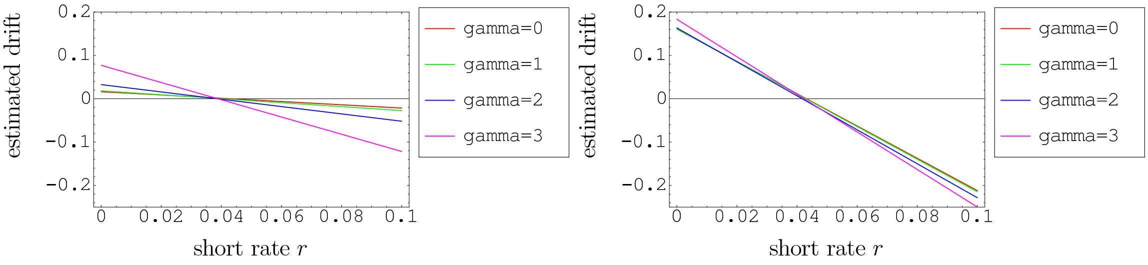

Figure 5.11: Estimated drifts $\alpha + \beta r$ (Euribor, 2007) for weights $w_{ij} = \tau_j^2$ (left) and $w_{ij} = 1/\tau_j^2$ (right).

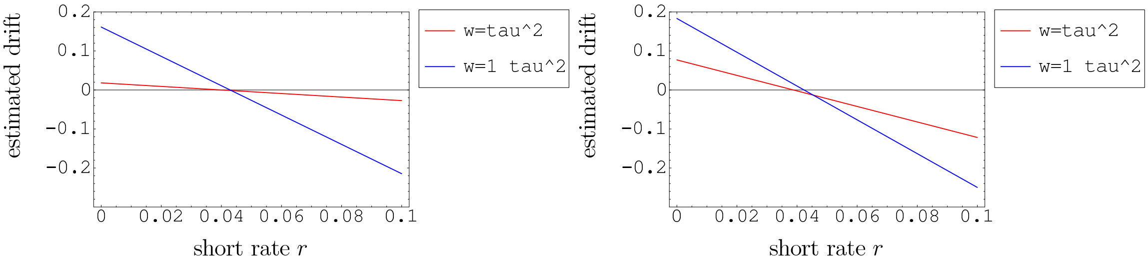

Figure 5.12: Estimated drifts $\alpha + \beta r$(Euribor, 2007) for $\gamma = 1$ (left) and $\gamma = 2$ (right).

5.11. Figure 5.12 compares drifts obtained for the same γ using these two weight functions. As it can be seen also from Table 5.3 and Table 5.4, weights $w = 1/\tau^2$ suggest stronger mean reversion, while the limit short rate (under the risk neutral measure) is similar in both cases.

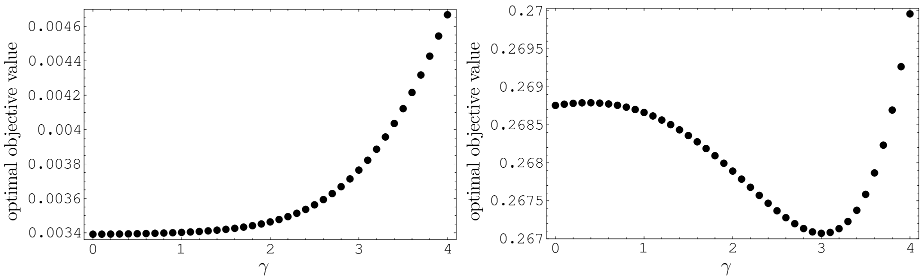

Figure 5.13: The objective function F as a function of γ for weights $w_{ij} = \tau_j^2$ (left) and $w_{ij} = 1/\tau_j^2$ (right).

Chapter 6

A survey of two-factor short rate models

Let us recall that in one-factor interest rate models, the term structure is completely determined by its origin, i.e. the short rate r. To allow different term structures, starting from the same short rate, another factor has to be introduced. It leads to a class of two-factor models. Interest rates are then functions not only of the short rate but they are also depending on the second factor of the model. Moreover, two-factor models allow for a larger variety shapes of term structures than one-factor models do. They can also predict market options data better, an analysis of option prices in one-factor and two-factor models was done in [13] and two-factor models were found to be more suitable. and accurate compared to one-factor models.

In this chapter we firstly derive the partial differential equation for the bond prices similarly as in the case of one-factor models. Then we describe several possible types of two-factor models arising from different choice of the second factor.

We consider a general two-factor models with the factors x, y, given by

$$\begin{aligned} dx &= \mu_x dt + \sigma_x dw_1, \\ dy &= \mu_y dt + \sigma_y dw_2, \end{aligned}$$

where correlation between dw_1 and dw_2 is a constant ρ, i.e. $E(dw_1 dw_2) = \rho dt$. The short rate is a function of these two factors, i.e. $r = r(x, y)$. Using the method in [29], we derive a PDE for bond prices in this model.

Denote by $P(x, y, t, T)$ the price of a zero coupon bond with maturity T, at the time t when the values of the factors are x and y. Then, by a multidimensional Itō's lemma (see for example [37], [29]) we get the stochastic differential equation for P:

$$dP = \mu dt + \sigma_1 dw_1 + \sigma_2 dw_2, \tag{6.1}$$

where $\mu = \mu(x, y, t, T)$, $\sigma_i = \sigma_i(x, y, t, T)$ are given by

$$\begin{aligned} \mu &= \frac{\partial P}{\partial t} + \mu_x \frac{\partial P}{\partial x} + \mu_y \frac{\partial P}{\partial y} + \frac{\sigma_x^2}{2}\frac{\partial^2 P}{\partial x^2} + \frac{\sigma_y^2}{2}\frac{\partial^2 P}{\partial y^2} + \rho\sigma_x\sigma_y \frac{\partial^2 P}{\partial x \partial y}, \\ \sigma_1 &= \sigma_x \frac{\partial P}{\partial x}, \\ \sigma_2 &= \sigma_y \frac{\partial P}{\partial y}. \end{aligned}$$

We construct a portfolio of bonds with three maturities T_1, T_2, T_3, consisting of V_1, V_2, V_3 units of the bonds. Change in its value π is then

$$\begin{aligned} d\pi &= V_1 dP(T_1) + V_2 dP(T_2) + V_3 dP(T_3) \\ &= (V_1\mu(T_1) + V_2\mu(T_2) + V_3\mu(T_3))\, dt \\ &\quad + (V_1\sigma_1(T_1) + V_2\sigma_1(T_2) + \sigma_1\mu(T_3))\, dw_1 \\ &\quad + (V_1\sigma_2(T_1) + V_2\sigma_2(T_2) + \sigma_2\mu(T_3))\, dw_2. \end{aligned} \tag{6.2}$$

By choosing V_1, V_2, V_3 so that

$$\begin{aligned} V_1\sigma_1(T_1) + V_2\sigma_1(T_2) + \sigma_1\mu(T_3) &= 0, \\ V_1\sigma_2(T_1) + V_2\sigma_2(T_2) + \sigma_2\mu(T_3) &= 0, \end{aligned}$$

the portfolio becomes deterministic

$$d\pi = V_1 dP(T_1) + V_2 dP(T_2) + V_3 dP(T_3)$$

and hence its return has to be equal to the riskless rate $r = r(x, y)$:

$$V_1\mu(T_1) + V_2\mu(T_2) + V_3\mu(T_3) = \pi r,$$

which can be written as

$$V_1(\mu(T_1) - r) + V_2(\mu(T_2) - r) + V_3(\mu(T_3) - r) = 0.$$

Hence we have a system of linear equations for the amounts V_1, V_2, V_3 of bonds in the portfolio:

$$\begin{pmatrix} \sigma_1(T_1) & \sigma_1(T_2) & \sigma_1(T_3) \\ \sigma_2(T_1) & \sigma_2(T_2) & \sigma_2(T_3) \\ \mu(T_1) - r & \mu(T_2) - r & \mu(T_3) - r \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This system has a nontrivial solution, if the row of the matrix are linearly dependent, i.e. some of them can be written as a linear combination of the preceding ones. If the second row is a multiple of the first one, substituting into (6.2) we only see one source of randomness and we essentially get a one-factor model. Hence the third row must be a combination of the previous two, i.e.

$$\mu(T_i) - r = \lambda_r \sigma_1(T_i)\sigma_1 + \lambda_2 \sigma_2(T_i).$$

Since T_i were arbitrary, this has to hold for any T_i and hence λ_1, λ_2 (so called market prices of risk of the factors) do not depend on the maturity of the bonds:

$$\lambda_1 = \lambda_1(x, y, t), \ \lambda_2 = \lambda_2(x, y, t).$$

Substituting values of μ, σ_1 and σ_2 we obtain the PDE satisfied by the bond price, which reads as

$$\begin{aligned} &\frac{\partial P}{\partial t} + (\mu_x - \lambda_1 \sigma_x)\frac{\partial P}{\partial x} + (\mu_y - \lambda_2 \sigma_y)\frac{\partial P}{\partial y} \\ +&\frac{\sigma_x^2}{2}\frac{\partial^2 P}{\partial x^2} + \frac{\sigma_y^2}{2}\frac{\partial^2 P}{\partial y^2} + \rho\sigma_x\sigma_y\frac{\partial^2 P}{\partial x \partial y} - r(x,y)P = 0. \end{aligned} \tag{6.3}$$

In the following sections we present and classify the two-factor models.

6.1 A stochastic parameter generalization of a one-factor model

Several two-factor models are simply obtained as a generalization of one-factor models, in which some parameter is now assumed to have a stochastic behavior. These models include:

- **A stochastic limit of the short rate.** A model with a stochastic short rate limit θ has been introduced in [8]. It is given by the following system of stochastic differential equations:

$$\begin{aligned} dr &= \kappa(\theta - r)dt + \sqrt{\sigma_0 + \sigma_1 r}dw_1 \\ d\theta &= \mu(\theta)dt + s(\theta)dw_2, \end{aligned}$$

  where the correlation $E(dw_1 dw_2)$ between the increments of the stochastic processes w_1 and w_2 is equal to ρdt. If we assume the market price of risk of the short rate to have a form $\lambda_1 + \lambda_2 r$ for some constants λ_1, λ_2, and the market price of risk being dependent only on θ, then the bond price has the form $P(\tau, r, \theta) = \exp(-A(\tau, \theta) - B(\tau)r)$. The function $B(\tau)$ can be found in the closed form. Moreover, if functions $\mu(\theta)$, $s(\theta)$, $l(\theta)$ are affine, then $A(\tau, \theta) = C(\tau) + \theta D(\tau)$ for some functions C, D. Hence interest rates are affine functions of r and θ (c.f. [8]).

- **A stochastic volatility.** In models with stochastic volatility, the second factor of the model is related to the volatility of the short rate. There are several possibilities for the choice of this factor. In the Fong-Vasicek model [24], it is given by

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}dw_1, \\ dy &= \kappa_2(\theta_2 - y)dt + v\sqrt{y}dw_2, \end{aligned}$$

  the short rate volatility is the square root of the second factor y. As before, increments of the processes can be correlated, their correlation is a constant ρ. If the market prices of risk are $\lambda_1\sqrt{y}$, resp. $\lambda_2\sqrt{y}$, then the bond price has the form $P(\tau, r, y) = A(\tau)e^{-B(\tau)r - C(\tau)y}$ (see [24]). It is possible to derive the closed form expressions for the function A, B and C, in terms of confluent hypergeometric functions with complex arguments (c.f. [39]).

  Anderson and Lund proposed several two-factor models in [4]. One of them is a model with a stochastic volatility given by

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sigma r^\gamma dw_1, \\ d\log\sigma^2 &= \kappa_2(\theta_2 - \log\sigma^2)dt + \xi dw_2, \end{aligned}$$

  where dw_1 and dw_2 are increments of independent Wiener processes.

Let us note that there are also three factor models, in which both these parameters are stochastic. The reader is referred to [16] for details.

6.2 A stochastic variable related to the short rate

Another popular class of models can be characterized by choosing the second factor to be a quantity observable on the market, which is assumed to be related to the short rate.

- **Consol rate.** In [11], the second factor is a yield ℓ of the consol bond. It is a bond with an infinite maturity paying a coupon. This bond is traded on the market, which allow to eliminate its market price of risk from the PDE for bond prices. Moreover, this equation does not contain the drift of the process for ℓ[1] A general form of this model is considered, for calibration purposes, the processes were specified as

$$\begin{aligned} dr &= (a_1 + b_1(\ell - r))dt + r\sigma_1 dw_1, \\ d\ell &= \ell(a_2 + b_2 r + c_2\ell)dt + \ell\sigma_2 dw_2 \end{aligned}$$

[1]Let us note that this is similar situation to Black-Scholes model for derivatives. The underlying asset is tradable, and in the PDE for derivatives prices there is neither drift of underlying, nor a market price of risk.

with dw_1 and dw_2 being independent increments of Wiener processes.

The model due to Schaefer and Schwarz [40] is a modification of the previous one. It models variables ℓ and $s = r - \ell$ (i.e. the difference between short rate and consol rate). Motivations for this transformation were empirical studies [5] and [6] confirming the independence of ℓ and s (opposed to r and ℓ). For the resulting PDE for bond prices, authors compute an approximate solution in the form $P(\tau, s, \ell) = X(\tau, s)Y(\tau, \ell)$ and compare it with the numerical solution.

Another approach how to model variables s and ℓ is given in [20], where GARCH models for these variables are used. They are based on the continuous model

$$\begin{aligned} d\ell &= (\alpha_{\ell_0} + \alpha_{\ell_1}\ell)dt + \sigma_\ell\, \ell^{\gamma_\ell} dw_1, & (6.4)\\ ds &= (\alpha_{s_0} + \alpha_{s_1}s)dt + \sigma_s\, \ell^{\gamma_s} dw_2, & (6.5)\end{aligned}$$

where increments dw_1 and dw_2 can be correlated, with a correlation coefficient ρdt. In contrary to results in [5] and [6], in all specifications of GARCH models, the correlation ρ is found to be statistically significantly different from zero. Parameter γ_s turns out to be zero, which means that volatility of s does not depend on its value. Parameter γ_ℓ is positive and in several models considered, it is close to $1/2$. From the deterministic parts of processes (6.4) and (6.5) follows mean reversion if $\alpha_{\ell_1} < 0$ and $\alpha_{s_1} < 0$, which was confirmed, although the speed of mean reversion is small.

- **European and domestic interest rates.** In [21] the interest rate in Spain (domestic rate r_d) before European monetary union was modelled. A process for the ECU interest rate (European rate r_e) is modelled by the one-factor Vasicek model:

  $$dr_e = c(d - r_e)dt + \sigma_e dw_e.$$

  A process for Spanish interest rate is modelled by

  $$dr_d = (a + b(r_e - r_d))dt + \sigma_2 dw_d,$$

  which can be written as

  $$dr_d = b\left(\frac{a}{b} + r_e - r_d\right)dt + \sigma_2 dw_2. \qquad (6.6)$$

  It can be seen from (6.6) that the domestic rate r_d is pushed to the value $\frac{a}{b} + r_e$. Wiener processes w_1, w_2 are assumed to be correlated, having a constant correlation ρdt of their increments dw_1, dw_2. Under the assumption of constant market prices of risks, European bond prices are given by a solution of the Vasicek model and domestic bond prices have the form

  $$P_d(\tau, r_d, r_e) = A(\tau)e^{-r_d B(\tau) - r_e C(\tau)},$$

with closed form expressions for A, B, C. Parameters of the model were estimated from short rate processes and market prices of risk from term structures. this model gives better performance when compared to a single factor Vasicek model.

A more general model of this kind can be found in [38]:

$$\begin{aligned} dr_e &= \mu_e(r_e)dt + \sigma_e(r_e)dw_1, \\ dr_d &= \mu_d(r_e - r_d)dt + \sigma_d(r_d)dw_2, \end{aligned}$$

where $\mu_e, \sigma_e, \mu_d, \sigma_d$ are general function. In [38] the authors estimated this model for Spanish and Italian data using nonparametric estimation methods.

6.3 Construction of the short rate from several processes.

In a generalized Cox-Ingersoll-Ross model [19], the short rate is assumed to be a sum of two independent Bessel square root processes, i.e.

$$\begin{aligned} r &= r_1 + r_2, \\ dr_1 &= \kappa_1(\theta_1 - r_1)dt + \sigma_1\sqrt{r_1}dw_1, \\ dr_2 &= \kappa_2(\theta_2 - r_2)dt + \sigma_2\sqrt{r_2}dw_2, \end{aligned} \tag{6.7}$$

where increments dw_1 and dw_2 are independent.

The economic theory, that was used in order to motivate this model, considers a model of economy with a single product. There are two derivations of the model presented in [19]. In the first one, the expected value and variance of the returns is characterized by two state variables. For a logarithmic utility function it can be derived that the real interest rate is a linear combination of state variables. After neglecting inflation and scaling the state variables, a model (6.7) is obtained. In the second derivation, the expected value and inflation is given just by one state variable. It determines the real interest rate. The second factor represents inflation. For a logarithmic utility function, the interest rate is again derived as a linear combination of the state variables. After their scaling, the model (6.7) is obtained. Details can be found in [19].

We made an assumption that market prices of risk of factors r_1 and r_2 are proportional to their square roots, i.e. $\lambda_i\sqrt{r_i}$, $i = 1, 2$, where λ_1, λ_2 are constants. In [19], this choice was also motivated by the economic theory. In this case, the bond price can be written as $P(\tau, r_1, r_2) = P_1(\tau, r_1)P_2(\tau, r_2)$, where PDEs for P_1 and P_2 are the equations arising from the one-factor CIR model. Hence they can be solved analytically.

In [31], the above model is formulated as a model with a stochastic volatility. The first factor is the short rate r, second factor V is its volatility. However, it is another kind of the stochastic model compared to those we considered in the previous sections. In this model, each of stochastic differential equations for r and V contains both Wiener processes. To formulate the model, we need to introduce the driving processes. Returns from the production are modelled by the stochastic differential equation

$$\frac{dQ}{Q} = (\mu X + \theta Y)dt + \sigma\sqrt{Y}dw_1,$$

where X, Y are mean reverting processes satisfying

$$\begin{aligned} dX &= \kappa_1(\theta_1 - X)dt + \sigma_1\sqrt{X}dw_2, \\ dY &= \kappa_2(\theta_2 - Y)dt + \sigma_2\sqrt{Y}dw_3. \end{aligned}$$

In [31] it was derived that the interest rate r is a linear combination of processes X and Y, i.e.

$$r = \alpha X + \beta Y,$$

where

$$\alpha = \mu\sigma_1^2,\ \beta = (\theta - \sigma^2)\sigma_2^2.$$

Then, the authors introduced constants γ, δ, η and ξ as follows:

$$\gamma = \frac{\kappa_1\theta_1}{\sigma_1^2},\ \delta = \kappa_1,\ \eta = \frac{\kappa_2\theta_2}{\sigma_2^2},\ \xi = \kappa_2.$$

They furthemore show that the processes for r and V can be written as

$$\begin{aligned} dr &= \left(\alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha}r - \frac{\xi - \delta}{\beta - \alpha}V\right)dt \\ &\quad + \alpha\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}dw_2 + \beta\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}dw_3, \\ dV &= \left(\alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta - \xi)}{\beta - \alpha}r - \frac{\beta\xi - \alpha\delta}{\beta - \alpha}V\right)dt \\ &\quad + \alpha^2\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}dw_2 + \beta^2\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}dw_3. \end{aligned}$$

This model was estimated by a discrete GARCH model, approximating the original continuous model. In their subsequent paper [30], authors used the GARCH method to estimate only a part of parameters and estimated the rest of them from the asymptotic distribution of r and V. In [28], a method of moments has been applied. In [10] the authors propose an iterative algorithm for estimation. A part of the parameters is estimated from the short rate data, the others are estimated from term structures. Results from one of these procedures is then used in the second one, which is repeated until the accuracy goal is achieved.

In a similar way, we also obtain a two-factor version of the Vasicek process where the short rate is a sum of two independent Ornstein-Uhlenbeck processes. There are also models having a deterministic function of time as one of the factors. For example the so called CIR++ model or Gaussian G++ model. Description of these models and their applications can be found, for example, in [12].

Chapter 7

The two-factor Vasicek Model

We remind ourselves that in the two-factor Vasicek model, the short rate is a sum of two independent Ornstein-Uhlenbeck processes r_1 and r_2:

$$\begin{aligned} dr_1 &= \kappa_1(\theta_1 - r_1)dt + \sigma_1 dw_1, \\ dr_2 &= \kappa_2(\theta_2 - r_2)dt + \sigma_2 dw_2. \end{aligned}$$

According to the general model (6.3), the bond price $P(\tau, r_1, r_2)$ is then a solution to the following partial differential equation of parabolic type

$$\begin{aligned} &-\frac{\partial P}{\partial \tau} + \left(\kappa_1(\theta_1 - r_1) - \tilde{\lambda}_1\sigma_1\right)\frac{\partial P}{\partial r_1} + \left(\kappa_2(\theta_2 - r_2) - \tilde{\lambda}_2\sigma_2\right)\frac{\partial P}{\partial r_2} \\ &+\frac{\sigma_1^2}{2}\frac{\partial^2 P}{\partial r_1^2} + \frac{\sigma_2^2}{2}\frac{\partial^2 P}{\partial r_2^2} - (r_1 + r_2)P = 0, \end{aligned} \tag{7.1}$$

which holds for any $r_1, r_2 \in (-\infty, \infty)$ and $\tau \in (0, \infty)$. A solution P satisfies the initial condition $P(0, r_1, r_2) = 1$ for each $r_1, r_2 \in (-\infty, \infty)$. Here, $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are market prices of risk, corresponding to the factors r_1 and r_2. If these functions are chosen to be constant, λ_1 and λ_2 resp., the solution of the resulting PDE has the form

$$P(\tau, r_1, r_2) = P_1(\tau, r_1)P_2(\tau, r_2), \tag{7.2}$$

where $P_i(\tau) = A_i(\tau)e^{-B_i(\tau)r}$, $i = 1, 2$, are bond prices in the Vasicek model with respective parameters given by (3.8) and (3.9). This property is derived directly by inserting (7.2) into (7.2).

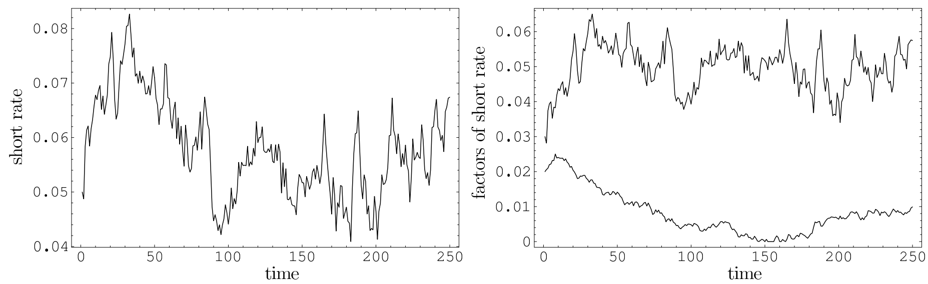

Figure 7.1: The short rate process r generated by the two-factor Vasicek model (left) and its components r_1 and r_2 (right).

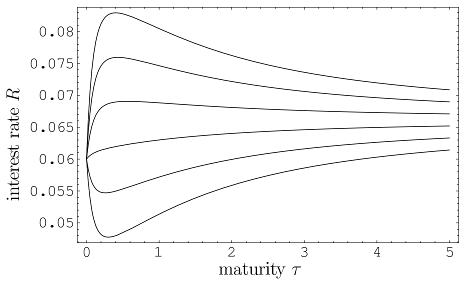

Figure 7.2: Term structures in the two-factor Vasicek model.

It follows from the form of a solution for bond prices (7.2) and the form of the bond price in the one-factor Vasicek model that the term structure in the two-factor Vasicek model is given by

$$R(\tau, r_1, r_2) = -\left(\frac{\ln A_1}{\tau} + \frac{\ln A_2}{\tau}\right) + \frac{B_1}{\tau} r_1 + \frac{B_2}{\tau} r_2. \tag{7.3}$$

Figures 7.1 and 7.2 show an example of a short rate process generated by the two-factor Vasicek model and term structures in this model.

7.1 Statistical properties of bond prices and interest rates

In practice, the components r_1 and r_2 of the short rate are not observable. An observable quantity is the short rate, i.e. their sum $r = r_1 + r_2$. Hence, we are interested in the conditional distribution of $P(\tau, r_1, r_2)$ and $R(\tau, r_1, r_2)$ under the constraint $r_1 + r_2 = r$, where the distributions of r_1 and r_2 are assumed to be their limiting distributions.

The limiting distributions of both Ornstein-Uhlenbeck processes forming the two-factor Vasicek model, are known to be normal distributions with density

$$f_i(x) = \frac{1}{\sqrt{2\pi\tilde{\sigma}_i^2}} e^{-\frac{(x-\theta_i)^2}{2\tilde{\sigma}_i^2}}, \quad i = 1, 2, \tag{7.4}$$

with $\tilde{\sigma}_i^2 = \frac{\sigma_i^2}{2\kappa_i}$ (c.f. [29]). Before deriving distributions of bond prices and interest rates, we formulate a theorem about conditional distribution of normal distributions. The proof can be found for example in [2].

Theorem 8. *[2, Theorem 4.12] Let X and Y be random variables with normal distributions $N(\mu_x, \sigma_x^2)$ and $N(\mu_y, \sigma_y^2)$ and let ρ be a correlation of X and Y. Then the conditional distribution of Y subject to $X = x$ is $N\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x - \mu_x), \sigma_y^2(1-\rho^2)\right)$.*

Theorem 9. *Consider the two-factor Vasicek model and the limiting distribution of the factors r_1 and r_2. Then:*

1. *The conditional density of the interest rate $R(\tau, r_1, r_2)$, subject to $r_1 + r_2 = r$, is given by*
$$f_R(x; \tau, r) = \frac{1}{\sqrt{2\pi\sigma_R^2}} e^{-\frac{(x-\mu_R)^2}{2\sigma_R^2}},$$
where
$$\begin{aligned}
\mu_R &= -\left(\frac{\ln A_1}{\tau} + \frac{\ln A_2}{\tau}\right) + \frac{1}{\tau}\left[B_1\theta_1 + B_2\theta_2 + \frac{B_1\tilde{\sigma}_1^2 + B_2\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}(r - (\theta_1 + \theta_2))\right], \\
\sigma_R^2 &= \frac{1}{\tau^2}\left(B_1^2\tilde{\sigma}_1^2 + B_2^2\tilde{\sigma}_2^2\right)\left(1 - \frac{(B_1\tilde{\sigma}_1^2 + B_2\tilde{\sigma}_2^2)^2}{(\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2)(B_1^2\tilde{\sigma}_1^2 + B_2^2\tilde{\sigma}_2^2)}\right).
\end{aligned}$$

2. *The conditional density of the bond price $P(\tau, r_1, r_2)$, subject to the condition $r_1 + r_2 = r$, is given by*
$$f_P(x) = \frac{1}{x}\frac{1}{\sqrt{2\pi\sigma_P^2}} e^{-\frac{(\ln x - \mu_P)^2}{2\sigma_P^2}}$$
for $x > 0$ and $f_P(x; \tau, r) = 0$ otherwise, where
$$\begin{aligned}
\mu_P &= \ln A_1 + \ln A_2 - \left((B_1\theta_1 + B_2\theta_2) + \frac{B_1\tilde{\sigma}_1^2 + B_2\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2 + \tilde{\sigma}_2}(r - (\theta_1 + \theta_2))\right), \\
\sigma_P^2 &= (B_1^2\tilde{\sigma}_1^2 + B_2^2\tilde{\sigma}_2^2)\left(1 - \frac{(B_1\tilde{\sigma}_1^2 + B_2\tilde{\sigma}_2^2)^2}{(\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2)(B_1^2\tilde{\sigma}_1^2 + B_2^2\tilde{\sigma}_2^2)}\right).
\end{aligned}$$

It means that the distribution of interest rates is a normal distribution $N(\mu_R, \sigma_R^2)$ and the distribution of bond prices is lognormal with the logarithm of a bond price having a normal distribution $N(\mu_P, \sigma_P^2)$.

Proof:

1. Since the term $-\left(\frac{\ln A_1}{\tau}+\frac{\ln A_2}{\tau}\right)$ in (7.3) is constant with respect to r_1, r_2, we will consider distribution of $\frac{B_1}{\tau}r_1+\frac{B_2}{\tau}r_2$, subject to the condition $r_1+r_2=r$. Define
$$X=r_1+r_2,\ Y=\frac{B_1}{\tau}r_1+\frac{B_2}{\tau}r_2,$$
then
$$X\sim N\left(\theta_1+\theta_2,\tilde{\sigma}_1^2+\tilde{\sigma}_2^2\right),\ Y\sim N\left(\frac{B_1}{\tau}\theta_1+\frac{B_2}{\tau}\theta_2,\left(\frac{B_1}{\tau}\right)^2\tilde{\sigma}_1^2+\left(\frac{B_2}{\tau}\right)^2\tilde{\sigma}_2^2\right),$$
$$Corr(X,Y)=\frac{\frac{B_1}{\tau}\tilde{\sigma}_1^2+\frac{B_2}{\tau}\tilde{\sigma}_2^2}{\sqrt{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\sqrt{\left(\frac{B_1}{\tau}\right)^2\tilde{\sigma}_1^2+\left(\frac{B_2}{\tau}\right)^2\tilde{\sigma}_2^2}}.$$
and the claim follows from the previous txheorem 8.

2. We have shown that $R\sim N(\mu_r,\sigma_R^2)$. Hence $-R\tau\sim N(-\mu_R\tau,\sigma_R^2\tau^2)$ and $P=e^{-R\tau}$ has a lognormal distribution with parameters given as in the theorem. For a density of a lognormal variable we refer to [2].

$\diamondsuit$

Figure 7.3 shows examples of the distributions. According to graphs shown in Figure 7.3, we expect that the variance of interest rates decreases for large maturities. We prove this property in the following theorem and give a condition guaranteeing a similar property also for the variance of bond prices.

Theorem 10. *Consider the limiting distribution of factors r_1 and r_2 given by (7.4). Then:*

1. *The conditional variance $Var(R(\tau,r_1,r_2|r_1+r_2=r)$ of interest rates (for a fixed r) converges to zero as time to maturity converges to infinity.*
2. *If*
$$\left(\theta_1-\frac{\sigma_1\lambda_1}{\kappa_1}-\frac{\sigma_1^2}{2\kappa_1^2}\right)+\left(\theta_2-\frac{\sigma_2\lambda_2}{\kappa_2}-\frac{\sigma_2^2}{2\kappa_2^2}\right)>0 \tag{7.5}$$
then the conditional variance $Var(P(\tau,r_1,r_2|r_1+r_2=r)$ of bond prices (for a fixed r) converges to zero as time to maturity converges to infinity.

Remark 5. *Recall that in the one-factor Vasicek model, in which $R(\tau,r)=-\frac{\ln A(\tau)}{\tau}+\frac{B(\tau)}{\tau}r$, we have*
$$\lim_{\tau\to\infty}R(\tau,r)=\lim_{\tau\to\infty}-\frac{\ln A(\tau)}{\tau}+\frac{B(\tau)}{\tau}r=\theta-\frac{\sigma\lambda}{\kappa}-\frac{\sigma^2}{2\kappa^2}$$

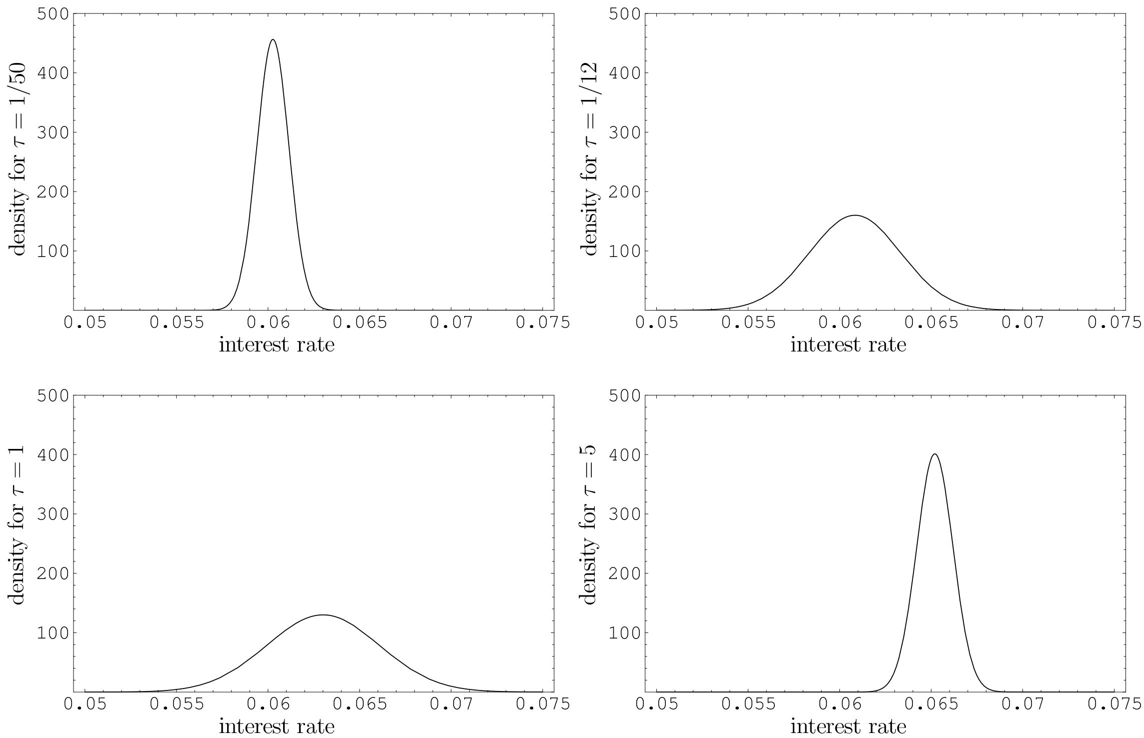

Figure 7.3: Distribution of interest rates in the two-factor Vasicek model.

(see (3.10)). In the two-factor Vasicek model the limit of term structures is

$$\begin{aligned}\lim_{\tau\to\infty} R(\tau,r_1,r_2) &= \lim_{\tau\to\infty} -\frac{\ln A_1(\tau)}{\tau} - \frac{\ln A_2(\tau)}{\tau} + \frac{B_1(\tau)}{\tau}r_1 + \frac{B_2(\tau)}{\tau}r_2 = \\ &= \left(\theta_1 - \frac{\sigma_1\lambda_1}{\kappa_1} - \frac{\sigma_1^2}{2\kappa_1^2}\right) + \left(\theta_2 - \frac{\sigma_2\lambda_2}{\kappa_2} - \frac{\sigma_2^2}{2\kappa_2^2}\right).\end{aligned}$$

Hence the condition (7.5) (which we will needed in the proof of Theorem 10) means that the limit of the term structures is positive.

Proof: We have already computed the variances, next we compute their limits.

1. In the previous section we derived variance of the interest rate R. Functions B_1 and B_2 have positive limits $1/\kappa_1$ and $1/\kappa_2$, hence we have

   $$Var(R(\tau,r_1,r_2)|r_1+r_2=r) = \frac{1}{\tau^2}\left(B_1^2\tilde{\sigma}_1^2 + B_2^2\tilde{\sigma}_2^2\right)\left(1 - \frac{(B_1\tilde{\sigma}_1^2 + B_2\tilde{\sigma}_2^2)^2}{(\tilde{\sigma}_1^2+\tilde{\sigma}_2^2)(B_1^2\tilde{\sigma}_1^2 + B_2^2\tilde{\sigma}_2^2)}\right),$$

   which converges to zero as $\tau \to \infty$.

2. Since the bond price has a lognormal distribution with $\ln P \sim N(\mu_P, \sigma_P^2)$, its conditional variance is (see [2])

   $$Var(P(\tau,r_1,r_2)|r_1+r_2=r) = e^{2\mu_P+\sigma_P^2}\left(e^{\sigma_P^2}-1\right).$$

   Since A_1, A_2 converge to zero if (7.5) is satisfied and B_1, B_2 have finite limits as $\tau \to \infty$, we obtain that $\mu_P \to -\infty$ and σ_P^2 has a finite limit as $\tau \to \infty$. It follows that the variance of P converges to zero.

$\diamondsuit$

7.2 Averaged values and confidence intervals

In this section we give averaged values of bond prices and interest rates for given short rate r, with respect to the conditional distribution of factor components r_1, r_2 of the short rate. Moreover we will analyze their confidence intervals.

In what follows, we will use the following notation for averaged bond prices and interest rates:

$$\begin{aligned}\tilde{P}(\tau,r) &= \langle P(\tau,r_1,r_2)|r_1+r_2=r\rangle, \\ \tilde{R}(\tau,r) &= \langle R(\tau,r_1,r_2)|r_1+r_2=r\rangle.\end{aligned}$$

Theorem 11. *[46] Averaged values of bond prices and interest rates, with respect to limit distributions of r_1, r_2, given that $r_1+r_2=r$, are*

1. $\tilde{R}(\tau, r) = -\left(\frac{\ln A_1}{\tau} + \frac{\ln A_2}{\tau}\right) + \frac{1}{\tau}\left[B_1\theta_1 + B_2\theta_2 + \frac{B_1\tilde{\sigma}_1^2 + B_2\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}(r - (\theta_1+\theta_2))\right]$,

2. $\tilde{P}(\tau, r) = \tilde{A}(\tau)e^{-\tilde{B}(\tau)r}$, *where*

$$\begin{aligned} \tilde{A}(\tau) &= A_1A_2\exp\left(-(B_1-B_2)\left(\theta_1 - (\theta_1+\theta_2)\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right) + \frac{1}{2}\frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}(B_1-B_2)^2\right), \\ \tilde{B}(\tau) &= \frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_1 + \frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_2. \end{aligned} \tag{7.6}$$

and $A_i = A_i(\tau)$, $B_i = B_i(\tau)$ *are given by (3.8) and (3.9)*

Proof: We have already computed the averaged interest rate as an expected value of interest rate distribution. The formula for the averaged bond price follows from lognormal distribution for bond prices. For the expected value of a lognormal variable we again refer the reader to [2]. $\diamondsuit$

It follows from Theorem 8 that the conditional distribution of r_1, subject to $r_1 + r_2 = r$, is normal $N(\mu_c, \sigma_c^2)$ with parameters

$$\mu_c = \theta_1 + \frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}(r - (\theta_1+\theta_2)),\ \sigma_c^2 = \frac{\tilde{\sigma}_1^2\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}.$$

Then $100p\%$ confidence interval (r_1^l, r_1^u) for r_1 can be constructed. We conclude that

$$\begin{aligned} P(\tau, r_1, r-r_1) &= A_1A_2e^{-B_2r-(B_1-B_2)r_1}, \\ R(\tau, r_1, r-r_1) &= -\left(\frac{\ln A_1}{\tau} + \frac{\ln A_2}{\tau}\right) + \left(\frac{B_1}{\tau} - \frac{B_2}{\tau}\right)r_1 + \frac{B_2}{\tau}r, \end{aligned}$$

are monotone functions of r_1 for fixed τ and r. Hence $P(\tau, r_1^l, r-r_1^l)$ and $P(\tau, r_1^u, r-r_1^u)$ are boundaries of the region, where the real bond price curve belongs to with a probability p. Similarly, $R(\tau, r_1^l, r-r_1^l)$ and $R(\tau, r_1^u, r-r_1^u)$ are boundaries of the confidence interval for term structures. Figure 7.4 shows averaged values and confidence intervals constructed in this way.

7.3 Relation of averaged values to one-factor models

Averaged values, computed in the previous section, are functions of time to maturity τ and the short rate r. It is a similar dependence as that of one-factor models. Therefore it is natural to study a question, whether there exists a one-factor interest rate model such that the averaged value $\tilde{P}(\tau, r)$ satisfies the corresponding PDE for bond prices. We restrict ourselves to interest rate models having the short rate r driven by the SDE:

$$dr = \mu(r)dt + \sigma(r)dw, \tag{7.7}$$

such that the drift μ, volatility σ and market price of risk λ are time independent.

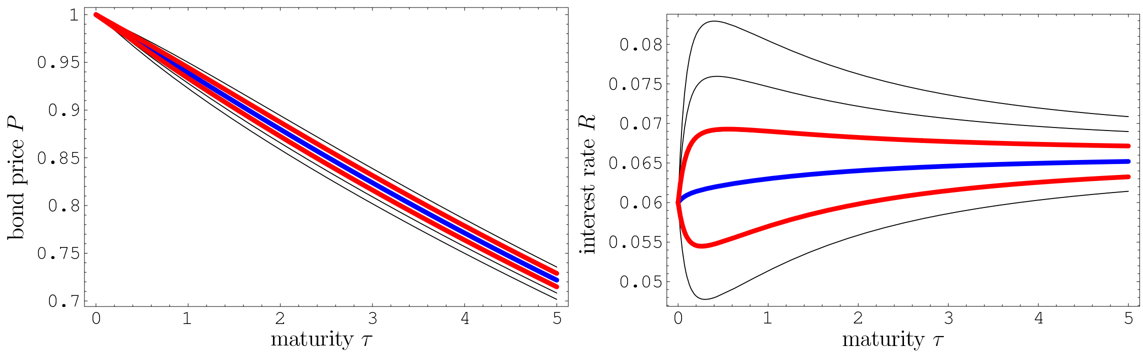

Figure 7.4: Averaged values (blue) and confidence intervals (red) for bond prices and interest rates in the two-factor Vasicek model.

Theorem 12. *[46, Theorem 3.1] Consider a class of one-factor models (7.7) where functions μ, σ, λ depend only on r and not on time t. Then there is no such a one-factor interest rate model, for which the averaged bond prices from the two-factor Vasicek model $\tilde{P}(\tau, r)$ satisfy the PDE*

$$-\frac{\partial P}{\partial \tau} + (\mu(r) - \lambda(r)\sigma(r))\frac{\partial P}{\partial r} + \frac{1}{2}\sigma(r)^2 \frac{\partial^2 P}{\partial r^2} - rP = 0 \tag{7.8}$$

for bond prices.

Proof: Suppose that the averaged bond price$\tilde{P}(\tau, r)$ is a solution of a one-factor model bond valuation PDE (7.8). Substituting it to this PDE yields

$$-\frac{\tilde{A}'(\tau)}{\tilde{A}(\tau)} + \tilde{B}'(\tau)r - (\mu(r) - \lambda(r)\sigma(r))\tilde{B}(\tau) + \frac{1}{2}\sigma^2(r)\tilde{B}^2(\tau) - r = 0. \tag{7.9}$$

It follows that $(\mu(r) - \lambda(r)\sigma(r))\tilde{B}(\tau) - \frac{1}{2}\sigma^2(r)\tilde{B}(\tau)^2$ is a linear function of r of the form

$$(\mu - \lambda\sigma)(r)\tilde{B}(\tau) - \frac{1}{2}\sigma^2(r)\tilde{B}(\tau)^2 = k_1(\tau) + k_2(\tau)r. \tag{7.10}$$

Moreover, we show that the following stronger condition has to be satisfied:

$$\sigma^2(r) = l_1 + l_2 r, \quad \text{where } l_2 \neq 0, \tag{7.11}$$

$$\mu(r) - \lambda(r)\sigma(r) = l_3 + l_4 r, \quad \text{where } l_4 \neq 0. \tag{7.12}$$

It means that the terms $\mu(r) - \lambda(r)\sigma(r)$ and $\sigma^2(r)$ do not contain nonlinear terms that could eventually vanish in $(\mu(r) - \lambda(r)\sigma(r))\tilde{B}(\tau) - \frac{1}{2}\sigma^2(r)\tilde{B}(\tau)^2$. Then we obtain the equation

$$\left(-\frac{\tilde{A}'(\tau)}{\tilde{A}(\tau)} - l_3\tilde{B}(\tau) + \frac{1}{2}l_1\tilde{B}^2(\tau)\right) + r\left(\tilde{B}'(\tau) - l_4\tilde{B}(\tau) + \frac{1}{2}l_2\tilde{B}(\tau) - 1\right) = 0.$$

Thus, the equation for $\tilde{B}$ reads as follows:

$$\begin{aligned}\tilde{B}'(\tau) &= 1-\left(\frac{1}{2}l_2-l_4\right)\tilde{B}(\tau),\\ \tilde{B}(0) &= 0.\end{aligned}$$

with $l_2, l_4 \neq 0$. This is an equation of the same form as the one appearing in the Cox-Ingersoll-Ross model and its solution is known in a closed form and it is given in the third chapter . However, the function

$$\tilde{B}(\tau)=\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_1(\tau)+\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_2(\tau)=c_0+c_1e^{-\kappa_1\tau}+c_2e^{-\kappa_2\tau}$$

for some constants c_0, c_1 and c_2, is not a function of this type.

To finish the proof, we prove (7.11) and (7.12). Firstly, we write the PDE in terms of $B_1(\tau)$ and $B_2(\tau)$ only, i.e.

$$\begin{aligned}-B_1(\tau)\left(\lambda_1\sigma_1-\kappa_1(\theta_1+\theta_2)\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right)-B_1(\tau)^2\left(\frac{1}{2}\sigma_1^2-\kappa_1\sigma_c^2\right)\\ -B_2(\tau)\left(\lambda_2\sigma_2-\kappa_2\theta_2-\kappa_2\left(\theta_1-(\theta_1+\theta_2)\frac{\sigma_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right)\right)-B_2^2(\tau)\left(\frac{1}{2}\sigma_2^2-\kappa_2\sigma_c^2\right)\\ -\left(\sigma_c^2(\kappa_1+\kappa_2)B_1(\tau)B_2(\tau)\right)+\left(-\kappa_1\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_1(\tau)-\kappa_2\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_2(\tau)\right)r\\ +\frac{1}{2}\sigma^2(r)\left(\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_1(\tau)+\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_2(\tau)\right)^2\\ -(\mu-\lambda\sigma)(r)\left(\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_1(\tau)+\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_2(\tau)\right) &= 0.\end{aligned}$$

The equality holds for all r and $\tau > 0$. Hence also the derivative of the left hand side with respect to τ is identically zero and its limit as $\tau \to 0^+$ is zero, too. This yields

$$\begin{aligned}&-\left[\left(\lambda_1\sigma_1-\kappa_1(\theta_1+\theta_2)\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right)+\left(\lambda_2\sigma_2-\kappa_2\theta_2-\kappa_2\left(\theta_1-(\theta_1+\theta_2)\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right)\right)\right]\\ &+\left[-\kappa_1\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}-\kappa_2\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right]r-(\mu(r)-\lambda(r)\sigma(r))=0.\end{aligned}$$

The proposition (7.11) follows, with

$$l_4=-\kappa_1\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}-\kappa_2\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}. \tag{7.13}$$

Hence $\sigma^2(r)$ is also a linear function of the form $l_1 + l_2 r$, as claimed in (7.12). What remains to show is that $l_2 \neq 0$. From (7.9) we see that the linear coefficient $k_2(\tau)$ of $(\mu(r)-\lambda(r)\sigma(r))\tilde{B}(\tau)-\frac{1}{2}\sigma^2(r)\tilde{B}(\tau)^2$ in (7.10) is given by

$$k_2(\tau)=\tilde{B}'(\tau)-1=-\kappa_1\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}-\kappa_2\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}. \tag{7.14}$$

From (7.13) we obtain that the linear coefficient in $(\mu(r)-\lambda(r)\sigma(r))\tilde{B}(\tau)$ is equal to

$$\left(-\kappa_1\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}-\kappa_2\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}\right)\left(\frac{\tilde{\sigma}_1^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_1(\tau)+\frac{\tilde{\sigma}_2^2}{\tilde{\sigma}_1^2+\tilde{\sigma}_2^2}B_2(\tau)\right).$$

But has a different form compared to (7.14). Hence, the linear coefficient in $\sigma^2(r)$ is not zero, which finishes the proof. $\diamondsuit$

Chapter 8

The two-factor Cox Ingersoll Ross model

In the two-factor CIR model, the short rate is assumed to be a sum of two independent Bessel square root processes r_1 and r_2:

$$\begin{aligned} dr_1 &= \kappa_1(\theta_1 - r_1)dt + \sigma_1\sqrt{r_1}dw_1, \\ dr_2 &= \kappa_2(\theta_2 - r_2)dt + \sigma_2\sqrt{r_2}dw_2. \end{aligned}$$

Hence, according to the general case (6.3), the bond price $P(\tau, r_1, r_2)$ is a solution to the following partial differential equation

$$\begin{aligned} &-\frac{\partial P}{\partial \tau} + \left(\kappa_1(\theta_1 - r_1) - \tilde{\lambda}_1\sigma_1\sqrt{r_1}\right)\frac{\partial P}{\partial r_1} + \left(\kappa_2(\theta_2 - r_2) - \tilde{\lambda}_2\sigma_2\sqrt{r_2}\right)\frac{\partial P}{\partial r_2} \\ &+\frac{\sigma^2}{2}r_1\frac{\partial^2 P}{\partial r_1^2} + \frac{\sigma_2^2}{2}r_2\frac{\partial^2 P}{\partial r_2^2} - (r_1 + r_2)P = 0, \end{aligned} \tag{8.1}$$

which holds for $r_1, r_2 \in (0, \infty)$ and $\tau \in (0, \infty)$, satisfying initial condition $P(0, r_1, r_2) = 1$ for all $r_1, r_2 > 0$. Here, $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ are market prices of risk, corresponding to each of the factors r_1 and r_2. If these functions are chosen to be proportional to $\sqrt{r_1}$ and $\sqrt{r_2}$, i.e. $\tilde{\lambda}_1 = \lambda_1\sqrt{r_1}$ and $\tilde{\lambda}_2 = \lambda_2\sqrt{r_2}$ for some constants λ_1 and λ_2, the solution of the resulting PDE has the form

$$P(\tau, r_1, r_2) = P_1(\tau, r_1)P_2(\tau, r_2), \tag{8.2}$$

where $P_i(\tau, r_i) = A_i(\tau)e^{-B_i(\tau)r_i}$, $i = 1, 2$, are bond prices in the CIR model with corresponding parameters (indexed by 1 and 2) given by (3.11). This can be shown by inserting (8.2) into (8.1).

To simplify the notation we define $A = A_1 A_2$. Then the bond price is given by

$$P(\tau, r_1, r_2) = A(\tau) e^{-B_1(\tau) r_1 - B_2(\tau) r_2} \tag{8.3}$$

and interest rate by

$$R(\tau, r_1, r_2) = -\frac{\ln A(\tau)}{\tau} + \frac{B_1(\tau)}{\tau} r_1 + \frac{B_2(\tau)}{\tau} r_2. \tag{8.4}$$

Typical decomposition of the short rate is similar as for the two-factor Vasicek model discussed in the previous chapter. Let us note that a simulation of trajectories can be performed using exact transition density which is a multiple of the noncentral χ^2 distribution. An algorithm for generating random numbers from this distribution can be found for example in [26].

8.1 Distribution of bond prices and interest rates

As in the case of a two-factor Vasicek model, we consider the limit distribution of r_1 and r_2. Distributions of the variables will then be considered with respect to these limit distributions and the condition on the observed short rate. The limiting distribution for CIR processes

$$dr_i = \kappa_i(\theta_i - r_i)dt + \sigma_i \sqrt{r_i} dw_i, \ i = 1, 2,$$

is known to be (see e.g. [29]) the gamma distribution $\Gamma(a_i, b_i)$ with parameters

$$a_i = \frac{2\kappa_i}{\sigma_i}, \ b_i = \frac{2\kappa_i}{\sigma_i}\theta_i$$

and the correspondning densities

$$f_i(x) = \frac{a_i^{b_i}}{\Gamma(b_i)} e^{-a_i x} x^{b_i - 1} \quad \text{for } x > 0 \tag{8.5}$$

and $f_i(x) = 0$ otherwise.

Theorem 13. *Let us consider the limiting gamma distribution of the two factors r_1 and r_2 given as in (8.5). Then:*

1. *The density function of the interest rate with maturity τ subject to the given level of short rate $r = r_1 + r_2$ is*

$$f_R(x) = \frac{g_1(\tilde{r}) g_2(r - \tilde{r})}{\int_0^r g_1(r_1) g_2(r - r_1) dr_1} \frac{1}{|B_2(\tau) - B_1(\tau)|}, \tag{8.6}$$

with

$$\tilde{r} = \frac{\tau x - (-\ln A(\tau) + B_2(\tau) r)}{B_1(\tau) - B_2(\tau)}$$

for x between values $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau} B_1(\tau) r$ and $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau} B_2(\tau) r$ and $f_R(x) = 0$ otherwise.

2. *The density function of the bond price with a maturity τ subject to the given level of short rate $r = r_1 + r_2$ is*

$$f_P(x) = \frac{1}{\tau x} f_R\left(-\frac{1}{\tau}\ln x\right), \tag{8.7}$$

where f_R is the density of interest rate given by (8.6).

Proof: The conditional density of r_1, subject to $r_1 + r_2 = r$, is

$$f(r_1, r) = \frac{f_1(r_1) f_2(r - r_1)}{\int_0^r f_1(s) f_2(r-s) ds}. \tag{8.8}$$

Now, we recall that the if $r_1 + r_2 = r$, the interest rate $R(\tau, r_1, r_2)$ can be written as $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r + \frac{1}{\tau}(B_1(\tau) - B_2(\tau))r_1$. Furthemore, we know that r_1 is from the interval $(0, r_1)$. Now, we consider two cases, depending on the sign of the term $B_1(\tau) - B_2(\tau)$.

Case 1: $B_1(\tau) - B_2(\tau) > 0$.
In this case, the minimal possible value of R is $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r$ and the maximal possible value is $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)r$. For x from the interval $(-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r, -\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)r)$, the distribution function of R, denoting the probability that $R < x$, is

$$F_R(x) = Prob(R < x) = \frac{\int_{\substack{r_1 \in (0,r) \\ R(\tau, r_1, r - r_1) < x}} g_1(r_1) g_2(r - r_1) dr_1}{\int_{r_1 \in (0,r)} g_1(r_1) g_2(r - r_1) dr_1}.$$

The condition $R(\tau, r_1, r - r_1) < x$ can be rewritten in the form $r_1 < \tilde{r}$, where

$$\tilde{r} = \frac{\tau x - (-\ln A(\tau) + B_2(\tau) r)}{B_1(\tau) - B_2(\tau)}. \tag{8.9}$$

Because $x \in (-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r,\ -\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)r)$, this quantity $\tilde{r}$ belongs to the interval $(0, r)$. Hence the distribution function F can be written as

$$F_R(x) = \frac{\int_0^{\tilde{r}} g_1(r_1) g_2(r - r_1) dr_1}{\int_0^r g_1(r_1) g_2(r - r_1) dr_1}$$

and we obtain the density by taking the derivative

$$\begin{aligned} f_R(x) &= F_R'(x) = \frac{g_1(\tilde{r}) g_2(r - \tilde{r})}{\int_0^r g_1(r_1) g_2(r - r_1) dr_1} \frac{d\tilde{r}}{dx} \\ &= \frac{g_1(\tilde{r}) g_2(r - \tilde{r})}{\int_0^r g_1(r_1) g_2(r - r_1) dr_1} \frac{\tau}{B_1(\tau) - B_2(\tau)}. \end{aligned}$$

Case 2: $B_1(\tau) - B_2(\tau) < 0$.
In this case, the minimal possible value of R is equal to $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)r$ and the maximal possible value is equal to $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r$. For x from the interval $(-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)r, -\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r)$ we compute the distribution function F in the same way as before:

$$F_R(x) = \frac{\int_0^{\tilde{r}} g_1(r_1)g_2(r-r_1)dr_1}{\int_0^r g_1(r_1)g_2(r-r_1)dr_1} = \frac{\int_{\tilde{r}}^r g_1(r_1)g_2(r-r_1)dr_1}{\int_0^r g_1(r_1)g_2(r-r_1)dr_1},$$

where

$$\tilde{r} = \frac{x\tau - (-\ln A(\tau) + B_2(\tau)r)}{B_1(\tau) - B_2(\tau)} \tag{8.10}$$

belongs to the interval $(0, r)$. Taking the derivative with respect to x we obtain the density function:

$$\begin{aligned} f_R(x) &= F_R'(x) = -\frac{g_1(\tilde{r})g_2(r-\tilde{r})}{\int_0^r g_1(r_1)g_2(r-r_1)dr_1}\frac{d\tilde{r}}{dx} \\ &= \frac{g_1(\tilde{r})g_2(r-\tilde{r})}{\int_0^r g_1(r_1)g_2(r-r_1)dr_1}\frac{\tau}{B_2(\tau) - B_1(\tau)}. \end{aligned} \tag{8.11}$$

Comparing (8.9) and (8.10) we see that, in the both cases, the quantity $\tilde{r}$ entering the formula for the density function is the same. $\diamondsuit$

Similarly as in the case of the two-factor Vasicek model, we are able to prove a theorem on the limit of variance for τ approaching infinity holds.

Theorem 14. *The conditional variance of both interest rates $R(\tau, r_1, r_2 | r_1 + r_2 = r)$ and bond prices $P(\tau, r_1, r_2 | r_1 + r_2 = r)$ converge to zero for a fixed r as $\tau \to \infty$.*

Proof: The proof is based on the intervals of possible values of interest rates and bond prices for different maturities. We use the following lemma.

Lemma 1. *Consider a family of the random variables X_τ, $\tau > 0$, with densities f_τ, means μ_τ and variances σ_τ^2. Suppose that X_τ takes values from the interval (a_τ, b_τ), where $(b_\tau - a_\tau) \to 0$ as $\tau \to \infty$. Then $\sigma_\tau^2 \to 0$ as $\tau \to \infty$.*

Proof of the lemma: Since X_τ takes values only in the interval (a_τ, b_τ), so does the mean value $\mu_\tau \in (a_\tau, b_\tau)$. The variance is given by

$$\sigma_\tau^2 = E\left((X_\tau - \mu_\tau)^2\right) = \int_{a_\tau}^{b_\tau} (x - \mu_\tau)^2 f_\tau(x)dx.$$

By the mean value theorem for integrals, we have

$$\sigma_\tau^2 = (\xi_\tau - \mu_\tau)^2 \int_{a_\tau}^{b_\tau} f_\tau(x)dx = (\xi_\tau - \mu_\tau)^2$$

for some $\xi_\tau \in (a_\tau, b_\tau)$. Since both ξ_τ and μ_τ belong to (a_τ, b_τ), we have

$$\sigma_\tau^2 = (\xi_\tau - \mu_\tau)^2 < (b_\tau - a_\tau)^2.$$

If $b_\tau - a_\tau$ converges to zero as $\tau \to \infty$, then the variance σ_τ^2 also converges to zero as $\tau \to \infty$. $\diamondsuit$

Proof of the theorem (continued): According to the previous lemma, it suffices to show that the lengths in the intervals of interest rate and bonds values converge to zero. Let us denote the interval of interest rates by (a_τ, b_τ). Then the interval of bond values is $(e^{-b_\tau \tau}, e^{-a_\tau \tau})$. We show that a_τ and b_τ have the same positive limit, from what it follows that both $b_\tau - a_\tau$ and $e^{-b_\tau \tau} - e^{a_\tau \tau}$ converge to zero.

Firstly, we note the that for functions $A_{cir}(\tau)$ and $B_{cir}(\tau)$ from 1-factor CIR model we have:

$$\lim_{\tau\to\infty} B_{cir}(\tau) = \lim_{\tau\to\infty} \frac{2(e^{\gamma\tau}-1)}{2\gamma + (\kappa+\lambda+\gamma(e^{\gamma\tau}-1)} = \frac{2}{\kappa+\lambda+\gamma},$$

because $\kappa + \lambda + \gamma = (\kappa+\lambda) + \sqrt{(\kappa+\lambda)^2 + 2\sigma^2} \neq 0$ for $\sigma \neq 0$. Using l'Hospital's rule we compute

$$\lim_{\tau\to\infty} \frac{1}{\tau} \log A_{cir}(\tau) = \lim_{\tau\to\infty} \frac{2\kappa\theta}{\sigma^2} \frac{1}{\tau} \log \left(\frac{2\gamma e^{\frac{1}{2}(\kappa+\lambda+\gamma)\tau}}{2\gamma + (\kappa+\lambda+\gamma)(e^{\gamma\theta}-1)} \right) = \frac{\kappa\theta}{\sigma^2}(\kappa+\lambda-\gamma).$$

Now, it follows that for 2-factor CIR model we have

$$\begin{aligned} \lim_{\tau\to\infty} -\frac{1}{\tau} \ln A(\tau) &= \lim_{\tau\to\infty} -\frac{1}{\tau} \ln A_1(\tau) - \frac{1}{\tau} \ln A_2(\tau) \\ = -\frac{\kappa_1\theta_1}{\sigma_1^2}(\kappa_1 + \lambda_1 - \gamma_1) &- \frac{\kappa_2\theta_2}{\sigma_2^2}(\kappa_2 + \lambda_2 - \gamma_2) > 0 \end{aligned}$$

and

$$\lim_{\tau\to\infty} \frac{1}{\tau} B_i(\tau) = 0, \quad \text{for } i = 1, 2.$$

Hence both $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)$ and $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)$ converge to same positive limit. Since a_τ and b_τ take values of $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_1(\tau)r$ and $-\frac{1}{\tau}\ln A(\tau) + \frac{1}{\tau}B_2(\tau)r$, our claim is proved. $\diamondsuit$

8.2 Averaged values and confidence intervals

In the following lemma we recall some of the properties of the Kummer confluent hypergeometric functions ${}_1F_1$. They will be used in the subsequent proof of Theorem 16.

Lemma 2.

1. *[1, formula (13.2.1)] The following equality holds:*
$$\int_0^r e^{-ax}x^{b-1}(r-x)^c dx = r^{b+c}\frac{\Gamma(b)\Gamma(1+c)}{\Gamma(1+b+c)}{}_1F_1(b,1+b+c,-ar)$$

2. *[1, formula (13.1.2)] The first terms in power series expansion of ${}_1F_1(a,b,z)$ are given by:*
$${}_1F_1(a,b,z) = 1+\frac{a}{b}z+\frac{a(a+1)}{b(b+1)}z^2+\frac{a(a+1)(a+2))}{b(b+1)(b+2)}z^3+\dots.$$

As in the case of the two-factor Vasicek model, we introduce the notation for the averaged bond price and interest rate:
$$\begin{aligned}\tilde{P}(\tau,r) &= \langle P(\tau,r_1,r_2)|r_1+r_2=r\rangle,\\ \tilde{R}(\tau,r) &= \langle R(\tau,r_1,r_2)|r_1+r_2=r\rangle.\end{aligned}$$
In the next theorem we explicitly compute these averaged value.

Theorem 15. *[45]*

1. *The averaged bond price with respect to limiting distributions of the processes r_1, r_2 given by (8.5) subject to $r_1+r_2=r$ is given by*
$$\tilde{P}(\tau,r) = Ae^{-Br}\,\frac{{}_1F_1(b_1,b_1+b_2,-((B_1-B_2)+(a_1-a_2)r))}{{}_1F_1(b_1,b_1+b_2,-(a_1-a_2)r)}, \tag{8.12}$$
where $A=A_1(\tau)A_2(\tau)$, $B_i=B_i(\tau)$, $i=1,2$, are given by (3.11).

2. *The averaged interest rate with respect to limiting distributions of the processes r_1, r_2 given by (8.5) subject to $r_1+r_2=r$ is given by*
$$\tilde{R}(\tau,r) = -\frac{\ln A}{\tau}+\frac{B_2}{\tau}r+\left(\frac{B_1}{\tau}-\frac{B_2}{\tau}\right)r\frac{b_1}{b_1+b_2}\frac{{}_1F_1(b_1,b_1+b_2+1,-(a_1+a_2)r)}{{}_1F_1(b_1,b_1+b_2,-(a_1+a_2)r)},$$
where $A=A_1(\tau)A_2(\tau)$, $B_i=B_i(\tau)$, $i=1,2$, are given by (3.11).

Proof: Firstly, we write the denominator appearing in the expression (8.8) for the density function $f(r_1,r)$ and the density itself in a form which will be useful later.
$$M(r) := \int_0^r f_1(r_1)f_2(r-r_1)dr_1 = \frac{a_1^{b_1}a_2^{b_2}}{\Gamma(b_1+b_2)}e^{-a_2 r}r^{b_1+b_2-1}{}_1F_1(b_1,b_1+b_2,-(a_1-a_2)r).$$
Substituting it into the density yields
$$\begin{aligned}f(r_1,r) &= \frac{1}{M(r)}f_1(r_1)f_2(r-r_1)\\ &= \frac{1}{{}_1F_1(b_1,b_1+b_2,-(a_1-a_2)r)}\frac{\Gamma(b_1+b_2)}{\Gamma(b_1)\Gamma(b_2)}\frac{e^{-(a_1-a_2)r_1}r_1^{b_1-1}(r-r_1)^{b_2-1}}{r^{b_1+b_2-1}}.\end{aligned} \tag{8.13}$$

Now, we can compute the expected values of bond prices and interest rates:

1. Substituting (8.13) into the expression for the averaged bond price gives

$$\begin{aligned}\tilde{P}(\tau,r) &= \int_0^r P(\tau,r_1,r-r_1)f(r_1,r)dr_1 \\ &= Ae^{-Br}\frac{{}_1F_1(b_1,b_1+b_2,-((B_1-B_2)+(a_1-a_2)r))}{{}_1F_1(b_1,b_1+b_2,-(a_1-a_2)r)}. \end{aligned} \tag{8.14}$$

2. Since

$$R(\tau,r_1,r_2|r_1+r_2=r) = -\frac{\ln A}{\tau}+\frac{B_2}{\tau}r+\left(\frac{B_1}{\tau}-\frac{B_2}{\tau}\right)r_1,$$

we need to compute the expected value of r_1. Substituting the density $f(r_1,r)$ yields

$$\begin{aligned}\langle r_1\rangle &= \int_0^r r_1\frac{f_1(r_1)f_2(r-r_1)}{\int_0^r f_1(s)f_2(r-s)ds}dr_1 \\ &= r\frac{b_1}{b_1+b_2}\frac{{}_1F_1(b_1,b_1+b_2+1,-(a_1+a_2)r)}{{}_1F_1(b_1,b_1+b_2,-(a_1+a_2)r)}.\end{aligned}$$

◇

Since for given r, the bond prices and interest rates are monotone functions of r_1, we can construct confidence intervals with the same methodology as in the case of the two-factor Vasicek model.

8.3 Relation of averaged values to one-factor models

Before stating the main theorem, concerned with averaged bond prices from the two-factor CIR model and one-factor models, we prove usefull properties of the averaging, that will be needed later.

Theorem 16. *[45, Theorem 3.1] Consider the averaged bond prices $\tilde{P}(\tau,r)$ from the previous section. They have the following properties:*

1. $\tilde{P}(\tau,r)\to A(\tau)$ *as* $r\to 0$,
2. $\frac{\partial\tilde{P}}{\partial\tau}(\tau,r)\to A'(\tau)$ *as* $r\to 0$,
3. $\frac{\partial\tilde{P}}{\partial r}(\tau,r)\to -A(\tau)\left(\frac{b_1}{b_1+b_2}B_1(\tau)+\frac{b_2}{b_1+b_2}B_2(\tau)\right)$ *as* $r\to 0$,
4. $\frac{\partial^2\tilde{P}}{\partial r^2}(\tau,r)$ *is bounded on the neighborhood of* $r=0$.

Proof:

1. Since both denominator and numerator of the fraction in (8.14) converge to unity as $r \to 0$, we have
$$\lim_{r\to 0} \tilde{P}(\tau, r) = A(\tau).$$

2. We compute the derivative of $\tilde{P}$ with respect to τ:
$$\begin{aligned} \frac{\partial \tilde{P}}{\partial \tau} &= \int_0^r \frac{\partial P}{\partial \tau}(\tau, r_1, r-r_1) f(r_1, r) dr_1 = \\ &= \left[\left(\frac{A'}{A} - B_2' r \right) - (B_1' - B_2') \frac{\int_0^r r_1 P(\tau, r_1, r-r_1) f(r_1, r) dr_1}{\int_0^r P(\tau, r_1, r-r_1) f(r_1, r) dr_1} \right] \tilde{P}. \end{aligned} \tag{8.15}$$
The numerator of the last fraction in (8.15) is positive for all $r > 0$ and can be bounded from above by $r \int_0^r P(\tau, r_1, r-r_1) f(r_1, r) dr_1$. Hence the fraction is positive and bounded from above by r, which implies that it converges to zero as $r \to 0$. Since we already know that $\tilde{P}(\tau, r) \to A(\tau)$ for $r \to 0$, we obtain from (8.15) that
$$\lim_{r\to 0} \frac{\partial \tilde{P}}{\partial \tau}(\tau, r) = A'(\tau).$$

3. By computing the derivative $\frac{\partial \tilde{P}}{\partial r}$ we obtain
$$\frac{\partial \tilde{P}}{\partial r} = \int_0^r \frac{\partial P}{\partial r}(\tau, r_1, r-r_1) f(r_1, r) + P(\tau, r_1, r-r_1) \frac{\partial f}{\partial r}(r_1, r) dr_1. \tag{8.16}$$
There are two derivatives that have to be computed: $\frac{\partial P}{\partial r}$ and $\frac{\partial f}{\partial r}$. Now, we evaluate these expressions. Firstly,
$$\frac{\partial P}{\partial r}(\tau, r_1, r-r_1) = -B_2(\tau) P(\tau, r_1, r-r_1). \tag{8.17}$$
Secondly,
$$\begin{aligned} \frac{\partial f}{\partial r}(r_1, r) &= \frac{f_1(r_1) f_2'(r-r_1)}{M(r)} - \frac{f_1(r_1) f_2(r-r_1)}{M^2(r)} M'(r) \\ &= f(r_1, r) \left[\frac{f_2'(r-r_1)}{f_2(r-r_1)} - \frac{\int_0^r f_1(s) f_2'(r-s) ds}{\int_0^r f_1(s) f_2(r-s) ds} \right] \end{aligned} \tag{8.18}$$
Noting that
$$\frac{f_2'(x)}{f_2(x)} = -a_2 + (b_2 - 1)\frac{1}{x}$$
and using it in (8.18) enables us to conclude
$$\frac{\partial f}{\partial r}(r_1, r) = f(r_1, r)(b_2 - 1) \left[\frac{1}{r - r_1} - \frac{\int_0^r \frac{1}{r-s} f_1(s) f_2(r-s) ds}{\int_0^r f_1(s) f_2(r-s) ds} \right]. \tag{8.19}$$

Substituting (8.17) and (8.19) into (8.16) yields after the rearrangement

$$\frac{\partial \tilde{P}}{\partial r} = \left[-B_2 + (b_2 - 1)\left(\frac{\int_0^r \frac{1}{r-r_1}\pi(\tau, r_1, r-r_1) f(r_1, r) dr_1}{\int_0^r \pi(\tau, r_1, r-r_1) f(r_1, r) dr_1} - \frac{\int_0^r \frac{1}{r-r_1} f_1(r_1) f_2(r-r_1) dr_1}{\int_0^r f_1(r_1) f_2(r-r_1) dr_1}\right)\right] \tilde{P}. \tag{8.20}$$

Let us denote

$$X_1 = \frac{\int_0^r \frac{1}{r-r_1}\pi(\tau, r_1, r-r_1) f(r_1, r) dr_1}{\int_0^r \pi(\tau, r_1, r-r_1) f(r_1, r) dr_1}, \quad X_2 = \frac{\int_0^r \frac{1}{r-r_1} f_1(r_1) f_2(r-r_1) dr_1}{\int_0^r f_1(r_1) f_2(r-r_1) dr_1}.$$

In this notation,

$$\frac{\partial \tilde{P}}{\partial r} = \left[-B_2 + (b_2 - 1)(X_1 - X_2)\right] \tilde{P}. \tag{8.21}$$

We write each of the expressions X_1 and X_2 in terms of the hypergeometric functions ${}_1F_1$:

$$X_1 = \frac{1}{r}\frac{b_1 + b_2 - 1}{b_2 - 1}\frac{{}_1F_1(b_1, b_1 + b_2 - 1, -((B_1 - B_2) + (a_1 - a_2)r))}{{}_1F_1(b_1, b_1 + b_2, -((B_1 - B_2) + (a_1 - a_2)r))} \tag{8.22}$$

and in a similar way

$$X_2 = \frac{1}{r}\frac{b_1 + b_2 - 1}{b_2 - 1}\frac{{}_1F_1(b_1, b_1 + b_2 - 1, -(a_1 - a_2)r)}{{}_1F_1(b_1, b_1 + b_2, -(a_1 - a_2)r)}. \tag{8.23}$$

Hence

$$X_1 - X_2 = \frac{1}{r}\frac{b_1 + b_2 - 1}{b_2 - 1}\left[\frac{G_1}{G_2} - \frac{G_3}{G_4}\right],$$

where we have denoted

$$\begin{aligned} G_1 &= {}_1F_1(b_1, b_1 + b_2 - 1, -((B_1 - B_2) + (a_1 - a_2))r), \\ G_2 &= {}_1F_1(b_1, b_1 + b_2, -((B_1 - B_2) + (a_1 - a_2))r), \\ G_3 &= {}_1F_1(b_1, b_1 + b_2 - 1, -(a_1 - a_2)r), \\ G_4 &= {}_1F_1(b_1, b_1 + b_2, -(a_1 - a_2)r). \end{aligned} \tag{8.24}$$

Because $G_2G_4 \to 1$ as $r \to 0$, we need to compute $G_1G_4 - G_2G_3$ in order to be able to compute the limit of (8.20). Since

$$\begin{aligned} G_1 &= 1 - \frac{b_1}{b_1 + b_2 - 1}((B_1 - B_2) + (a_1 - a_2))r + o(r), \\ G_2 &= 1 - \frac{b_1}{b_1 + b_2}((B_1 - B_2) + (a_1 - a_2))r + o(r), \\ G_3 &= 1 - \frac{b_1}{b_1 + b_2 - 1}(a_1 - a_2)r + o(r), \\ G_4 &= 1 - \frac{b_1}{b_1 + b_2}(a_1 - a_2)r + o(r), \end{aligned} \tag{8.25}$$

as $r \to 0$, we have

$$G_1G_4 - G_2G_3 \;=\; r\left(-\frac{b_1}{b_1+b_2-1}+\frac{b_1}{b_1+b_2}\right)+o(r) \tag{8.26}$$

as $r \to 0$. Hence

$$X_1 - X_2 \;=\; \frac{b_1+b_2-1}{b_2-1}\frac{1}{G_2G_4}\left[(B_1-B_2)\left(-\frac{b_1}{b_1+b_2-1}+\frac{b_1}{b_1+b_2}\right)+\frac{o(r)}{r}\right]$$

and

$$\lim_{r\to 0} X_1 - X_2 = \frac{b_1+b_2-1}{b_2-1}(B_1-B_2)\left(-\frac{b_1}{b_1+b_2-1}+\frac{b_1}{b_1+b_2}\right).$$

Finally, we can compute the limit of (8.20)

$$\begin{aligned}\lim_{r\to 0}\frac{\partial \tilde{P}}{\partial r}(\tau,r) &= \lim_{r\to 0}\left[-B_2+(b_2-1)(X_1-X_2)\right]\tilde{P}\\ &= A\left[-B_2+(b_1+b_2-1)(B_1-B_2)\left(-\frac{b_1}{b_1+b_2-1}+\frac{b_1}{b_1+b_2}\right)\right]\\ &= -A\left[\frac{b_1}{b_1+b_2}B_1+\frac{b_2}{b_1+b_2}B_2\right].\end{aligned}$$

4. We show that there is a finite limit of $\frac{\partial^2\tilde{P}}{\partial r^2}(\tau,r)$ as $r\to 0$, from which the boundedness of $\frac{\partial^2\tilde{P}}{\partial r^2}$ follows.

   According to (8.20) we have

$$\frac{\partial^2\tilde{P}}{\partial r^2} = \frac{\partial\tilde{P}}{\partial r}\left[-B_2+(b_2-1)(X_1-X_2)\right]+\tilde{P}\frac{\partial\left[-B_2+(b_2-1)(X_1-X_2)\right]}{\partial r}.$$

   From the definition of X_1 and X_2 and already computed limits it follows that it suffices to show the existence of the finite limit of $\frac{\partial}{\partial r}\left(\frac{1}{r}F(r)\right)$ for $r\to 0+$, where

$$F(r) = \frac{G_1(r)}{G_2(r)} - \frac{G_3(r)}{G_4(r)}. \tag{8.27}$$

   Assuming $F(r)$ has the power series expansion $F(r)=\sum_{k=0}^{\infty}a_kr^k$, the condition $a_0=0$ is sufficient for boundedness of the term $\frac{\partial}{\partial r}\left(\frac{1}{r}F(r)\right)$ in the neighborhood of $r=0$, which holds for (8.27). ♢

Now we state the main result on the nonexistence of a one-factor model describing the averaged bond price $\tilde{P}$. It has been shown by the author in [45].

Theorem 17. *[45, Theorem 3.3] Consider averaged bond prices $\tilde{P}(\tau, r)$ obtained from the two-factor CIR model and a class of one-factor short rate models*

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

satisfying

1. *functions μ, σ from the short rate process and market price of risk λ depend only on r and not on t,*
2. *functions μ, σ, λ are continuous at $r = 0$, $\sigma(0) = 0$.*
3. *volatility parameters of the factors from the two-factor CIR model are mutually different, i.e. $\sigma_1 \neq \sigma_2$.*

Then, there is no such a one-factor interest rate model, for which the averaged bond price $\tilde{P}(\tau, r)$ satisfies the PDE for bond prices

$$-\frac{\partial P}{\partial \tau} + (\mu(r) - \lambda(r)\sigma(r))\frac{\partial P}{\partial r} + \frac{1}{2}\sigma(r)^2\frac{\partial^2 P}{\partial r^2} - rP = 0 \tag{8.28}$$

for all $r \geq 0$, $\tau > 0$.

Proof: By taking the limit $r \to 0$ in the PDE (8.28), using the results from the previous theorem, we obtain for all $\tau > 0$:

$$-A'(\tau) + \mu(0^+)(-A(\tau))\left(\frac{b_1}{b_1 + b_2}B_1(\tau) + \frac{b_2}{b_1 + b_2}B_2(\tau)\right) = 0.$$

From this we calculate the value of the function μ for $r = 0$:

$$\mu(0^+) = -\frac{A'(\tau)}{A(\tau)}\frac{1}{\frac{b_1 B_1(\tau)}{b_1+b_2} + \frac{b_2 B_2(\tau)}{b_1+b_2}} = -\frac{A'(\tau)}{A(\tau)}\frac{b_1 + b_2}{b_1 B_1(\tau) + b_2 B_2(\tau)}.$$

It follows that

$$-\frac{A'(\tau)}{A(\tau)}\frac{b_1 + b_2}{b_1 B_1(\tau) + b_2 B_2(\tau)} = K_1, \tag{8.29}$$

for all $\tau > 0$ where K_1 is a constant independent of τ.

Now we recall that the function $A(\tau)$ from the two-factor CIR model can be written as $A(\tau) = A_1(\tau)A_2(\tau)$, where $A_1(\tau)$ and $A_2(\tau)$ are functions appearing in the original CIR model, corresponding to each of the equations for P_1 and P_2, resp. Hence they satisfy

$$A_i'(\tau) = -\kappa_i\theta_i A_i(\tau)B_i(\tau) \quad i = 1, 2.$$

Therefore

$$\frac{A'(\tau)}{A(\tau)} = \frac{A_1'(\tau)A_2(\tau) + A_1(\tau)A_2'(\tau)}{A_1(\tau)A_2(\tau)} = \frac{A_1'(\tau)}{A_1(\tau)} + \frac{A_2'(\tau)}{A_2(\tau)} = -\kappa_1\theta_1 B_1(\tau) - \kappa_2\theta_2 B_2(\tau).$$

Thus the equality (8.29) can be rewritten as

$$K_1 = -\frac{A'(\tau)}{A(\tau)}\frac{b_1+b_2}{b_1B_1(\tau)+b_2B_2(\tau)} = (\kappa_1\theta_1B_1(\tau)+\kappa_2\theta_2B_2(\tau))\frac{b_1+b_2}{b_1B_1(\tau)+b_2B_2(\tau)}.$$

Since b_1+b_2 is constant, the only important part is the following fraction:

$$\frac{\kappa_1\theta_1B_1(\tau)+\kappa_2\theta_2B_2(\tau)}{b_1B_1(\tau)+b_2B_2(\tau)} = K,$$

which has to be equal to some constant K. It implies that

$$\kappa_1\theta_1B_1(\tau)+\kappa_2\theta_2B_2(\tau) = K(b_1B_1(\tau)+b_2B_2(\tau))$$

and so

$$(\kappa_1\theta_1 - Kb_1)B_1(\tau) = (Kb_2 - \kappa_2\theta_2)B_2(\tau)$$

for each $\tau > 0$. It is possible in two ways:

1. $\kappa_1\theta_1 - Kb_1 = 0$, $Kb_2 - \kappa_2\theta_2 = 0$,
2. $B_1(\tau) = cB_2(\tau)$, where c is a constant.

Now we look at each of these possibilities:

1. The same constant K appears in both equalities. From the first one (i.e. $\kappa_1\theta_1 - Kb_1 = 0$), we get $K = \frac{\kappa_1\theta_1}{b_1}$ and by substituting the value of $b_1 = \frac{2\kappa_1\theta_1}{\sigma_1^2}$, we obtain $K = \frac{\sigma_1^2}{2}$. In the same way, from the second equality (i.e. $Kb_2 - \kappa_2\theta_2 = 0$), we obtain $K = \frac{\sigma_2^2}{2}$. But by the hypothesis, $\sigma_1^2 \neq \sigma_2^2$, which is a contradiction.

2. We recall the equation for B_1 from the CIR model:

$$-B_1'(\tau) = (\kappa_1 + \lambda_1\sigma_1)B_1(\tau) + \frac{1}{2}\sigma_1^2B_1(\tau)^2 - 1. \tag{8.30}$$

   An analogous equation for $B_2(\tau)$ yields

$$-B_2'(\tau) = (\kappa_2 + \lambda_2\sigma_2)B_2(\tau) + \frac{1}{2}\sigma_2^2B_2(\tau)^2 - 1. \tag{8.31}$$

   Since $B_1(\tau) = cB_2(\tau)$, we obtain another expression for B_1:

$$-B_1'(\tau) = c\left[(\kappa_2 + \lambda_2\sigma_2)B_2(\tau) + \frac{1}{2}\sigma^2B_2(\tau)^2 - 1\right]. \tag{8.32}$$

   The right-hand sides of (8.30) and (8.32) must equal to:

$$c\left[(\kappa_2 + \lambda_2\sigma_2)B_2(\tau) + \frac{1}{2}\sigma^2B_2(\tau)^2 - 1\right] = (\kappa_1 + \lambda_1\sigma_1)B_1(\tau) + \frac{1}{2}\sigma_1^2B_1(\tau)^2 - 1$$

for all $\tau > 0$. By continuity, the equality holds also in the limit $\tau = 0^+$. From this, we get $c = 1$ and hence the functions $B_1(\tau)$ and $B_2(\tau)$ coincide. We denote this function by $B(\tau)$. By subtracting equations (8.30) and (8.31) we obtain:

$$[-(\kappa_1 + \lambda_1\sigma_1) + (\kappa_2 + \lambda_1\sigma_1)]\, B(\tau) + \left[-\frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_2^2\right] B^2(\tau) = 0$$

and, dividing by a nonzero term $B(\tau)$ we obtain

$$[-(\kappa_1 + \lambda_1\sigma_1) + (\kappa_2 + \lambda_1\sigma_1)] - \frac{1}{2}\left[\sigma_2^2 - \sigma_1^2\right] B(\tau) = 0.$$

Since $\sigma_1 \neq \sigma_2$, it implies that $B(\tau)$ is a constant function, which is an obvious contradiction.

Since both possibilities lead to a contradiction, the theorem is proved. $\diamondsuit$

Chapter 9

The Fong-Vasicek model with stochastic volatility

In the Fong-Vasicek model the stochastic process driving the short rate is given by the following system of stochastic differential equations:

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}dw_1, \\ dy &= \kappa_2(\theta_2 - y)dt + v\sqrt{y}dw_2, \end{aligned} \tag{9.1}$$

where the correlation $E(dw_1 dw_2)$ of differentials dw_1 and dw_2 of Wiener processes is ρdt. If the market prices of risk are $\lambda_1\sqrt{y}$, resp. $\lambda_2\sqrt{y}$, then the PDE for the bond prices 6.3 reads as follows:

$$\begin{aligned} -\frac{\partial P}{\partial \tau} + (\kappa_1(\theta_1 - r) - \lambda_1 y)\frac{\partial P}{\partial r} + (\kappa_2(\theta_2 - y) - \lambda_2 v y)\frac{\partial P}{\partial y} \\ + \frac{y}{2}\frac{\partial^2 P}{\partial r^2} + \frac{v^2 y}{2}\frac{\partial^2 P}{\partial y^2} + \rho v y \frac{\partial^2 P}{\partial r \partial y} - rP = 0. \end{aligned} \tag{9.2}$$

It is known that a solution to (9.2) has the form $P(\tau, r, y) = A(\tau)e^{-B(\tau)r - C(\tau)y}$ (see [24]). As noted in Chapter 6, there are several possibilities to characterize the functions A, B, C. We will use a system of ordinary differential equations for these functions because it will be helpful in deriving the properties of the Fong-Vasicek model.

9.1 Qualitative properties of bond prices and term structures

In the following theorem we give the characterization of the bond price by a system of ordinary differential equations. This form will be used later when proving the properties of the model.

Theorem 18. *[42] A solution of the PDE for bond prices (9.2) has the form*

$$P(\tau, r, y) = A(\tau) e^{-B(\tau) r - C(\tau) y}, \tag{9.3}$$

for $r > 0$, $y > 0$ and $\tau > 0$, where functions $A = A(\tau)$, $B = B(\tau)$, $C = C(\tau)$ satisfy the following system of ordinary differential equations:

$$\begin{aligned}
A' &= -A\left(\kappa_1\theta_1 B + \kappa_2\theta_2 C\right), \\
B' &= -\kappa_1 B + 1, \\
C' &= -\lambda_1 B - \kappa_2 C - \lambda_2 v C - \frac{B^2}{2} - \frac{v^2 C^2}{2} - v\rho B C,
\end{aligned} \tag{9.4}$$

with initial conditions $A(0) = 1$, $B(0) = 0$, $C(0) = 0$. This can be represented in the following form:

$$B = \frac{1}{\kappa_1}\left(1 - e^{-\kappa_1 \tau}\right), \tag{9.5}$$

$$C' = -\lambda_1 B - \frac{B^2}{2} - \left(\kappa_2 + \lambda_2 v + v\rho B\right) C - \frac{v^2}{2} C^2, \ C(0) = 0, \tag{9.6}$$

$$A = \exp\left(-\theta_1 \tau + \theta_1 B - \kappa_2\theta_2 \int_0^\tau C(s) ds\right). \tag{9.7}$$

Proof: The assertions of the theorem follow after inserting the form of the solution (9.3) into the PDE (9.2). An equation for $B(\tau)$ can be solved analytically. The result is then substituted into the equation for $C(\tau)$, which we can solve numerically by the Runge-Kutta method. Finally, we integrate the equation for $A(\tau)$ and use the results for functions B and C. ◇

Corollary 2. *Interest rates are linear in both short rate r and volatility y and they are given by*

$$R(\tau, r, y) = -\frac{\ln A(\tau)}{\tau} + \frac{B(\tau)}{\tau} r + \frac{C(\tau)}{\tau} y.$$

In what follows, we will assume that the structural condition

$$\lambda_1 \leq -\frac{1}{2\kappa_1} \tag{9.8}$$

is satisfied. We will need this assumption in order to prove qualitative properties of the solution P given by (9.3).

In the next theorem we prove some of the properties of the functions A, B, C appearing in the definition of the function P.

Theorem 19. *[42] Under the assumption (9.8), the following conditions hold:*

1. $C'(0) = 0$, $C''(0) = -\lambda_1$,
2. *For every* $\tau > 0$: $0 < A(\tau) < 1$. $B(\tau) > 0$, $C(\tau) > 0$,
3. $A(\tau) \to 0$ *for* $\tau \to \infty$,
4. $C(\tau)$ *is bounded on* $[0, \infty)$.

Proof:

1. It follows from the differentiating the ODE for $C(\tau)$ and using the continuity of C and its derivatives at $\tau = 0$.

2. From the previous statement and the assumption $\lambda_1 \leq -1/2\kappa_1 < 0$ it follows that $C(\tau) > 0$ on some neighborhood of $\tau = 0$. Hence it suffices to show that $C'(\tau) > 0$ whenever $C(\tau) = 0$. To prove the above claim, we write $C'(\tau)$ in the following form:

$$C'(\tau) = -\lambda_1 B(\tau) - \frac{B^2(\tau)}{2} = -\frac{1 - e^{-\kappa_1 \tau}}{2\kappa_1^2} \left(2\lambda_1 \kappa_1 + 1 - e^{-\kappa_1 \tau}\right) > 0,$$

   provided $C(\tau) = 0$ and $\lambda_1 \leq -1/2\kappa_1$.

   Positiveness of $B(\tau)$ follows directly from the expression of this function (notice that $B(\tau) < \tau$.

   The function $A(\tau)$ is positive. Its upper bound follows from the following estimate:

$$\begin{aligned} A(\tau) &= \exp\left(-\theta_1(\tau - B(\tau)) - \kappa_2\theta_2 \int_0^\tau C(s)ds\right) \\ &< \exp\left(-\theta_1(\tau - B(\tau))\right) < 1. \end{aligned}$$

   The first inequality follows from the positiveness of C and the second one from the positiveness of the difference $\tau - B(\tau)$.

3. We have already shown that

$$0 < A(\tau) < \exp\left(-\theta(\tau - B(\tau))\right).$$

   Since $\tau - B(\tau)$ converges to infinity as $\tau \to \infty$, we obtain that $A(\tau) \to 0$ as $\tau \to \infty$.

4. It suffices to show that there exists a constant $K > 0$ such that $C'(\tau) < 0$ whenever $C(\tau) = K$. Notice that $B(\tau) < 1/\kappa_1$. Since $-v\rho B < 0$ for $\rho < 0$ and

$-v\rho B < \frac{\rho v}{\kappa_1}$ for $\rho > 0$, we have $-v\rho B < \min\left(0, -\frac{\rho v}{\kappa_1}\right)$. Using this inequality and the assumption (9.8), from (9.6) we obtain the estimate

$$C'(\tau) < -\frac{\lambda_1}{\kappa_1} + \left(-\kappa_2 - \lambda_2 v + \min\left(0, -\frac{\rho v}{\kappa_1}\right)\right) K - \frac{v^2}{2} K^2,$$

which is satisfied for any τ such that $C(\tau) = K$. Taking K sufficiently large, the proof of the statement 4 follows.

$\diamondsuit$

Remark 6. *Note that if $\lambda_1 > 0$, then $C''(0) = -\lambda_1 < 0$. Since $C(0) = 0$ and $C'(0) = 0$, we thus have that $C(\tau) < 0$ for small $\tau > 0$. Hence, for small times to maturity τ, the interest rates $R(\tau, r, y)$ becomes negative for large volatility y. This is avoided by assuming (9.8). It is a stronger condition, but we needed it in this form to prove the assertions of the theorem.*

9.2 Distribution of stochastic bond prices and interest rates

Let us recall that $P(\tau, r, y)$ is the price of a bond maturing at time τ for a given values of the short rate r and volatility y. Unlike the short rate r, the volatility y is not an observable variable in the real market. It suggests investigation of $P(\tau, r, y)$ for the given τ and r as a function of the random variable y.

In what follows, we will assume that the value of the short rate r at time to maturity τ is known from the market data. The hidden parameter in the model is the volatility y which is supposed to be driven by a Bessel square root process. We already know from the chapter 8 on two-factor CIR model, that its limiting density f_y is a density of the gamma distribution $\Gamma(\beta, \alpha)$ with shape parameters $\beta = \frac{2\kappa_2}{v^2}$, $\alpha = \frac{2\kappa_2}{v^2}\theta_2$, i.e.

$$f_y(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta x} x^{\alpha - 1} \tag{9.9}$$

for $x > 0$ and $f_y(x) = 0$ otherwise. It enables us to compute the distribution of $P(\tau, r, y)$ and $R(\tau, r, y)$ now.

Theorem 20. *Under the assumption (9.8) the density functions of the bond prices $P(\tau, r, y)$ and interest rates $R(\tau, r, y)$ with respect to the limiting distribution (9.9) of the process y are given by*

$$f_P(x) = f_y\left(-\frac{B(\tau)}{C(\tau)} r - \frac{1}{C(\tau)} \ln \frac{x}{A(\tau)}\right) \frac{1}{C(\tau) x} \tag{9.10}$$

and

$$f_R(x) = f_y\left(\frac{1}{C(\tau)} \left(\tau x + \ln A(\tau) - B(\tau)\right)\right) \frac{\tau}{C(\tau)}, \tag{9.11}$$

where f_y is density of the limiting gamma distribution.

Proof: First, we compute distribution functions and then by differentiating them we obtain densities.

Since $P(\tau, r, y)$ is a decreasing function of y, the range of its possible values is the interval $\left(0, A(\tau)e^{-B(\tau)r}\right]$. Hence outside of this interval, the density is vanishing. For $x \in \left(0, A(\tau)e^{-B(\tau)r}\right]$ we have the following expression for the cumulative distribution function $F_P(x)$:

$$\begin{aligned} F_P(x) &= \text{Prob}\left[A(\tau)e^{-B(\tau)r - C(\tau)y} < x\right] = \text{Prob}\left[y > -\frac{B(\tau)}{C(\tau)}r - \frac{1}{C(\tau)}\ln\frac{x}{A(\tau)}\right] \\ &= 1 - F_y\left(-\frac{B(\tau)}{C(\tau)}r - \frac{1}{C(\tau)}\ln\frac{x}{A(\tau)}\right), \end{aligned}$$

where we have used the positiveness of $C(\tau)$, and so

$$f_P(x) = F_P'(x) = f_y\left(-\frac{B(\tau)}{C(\tau)}r - \frac{1}{C(\tau)}\ln\frac{x}{A(\tau)}\right)\frac{1}{C(\tau)x}.$$

Similarly, because of increasing dependence of $R(\tau, r, y)$ on y, the range of possible values for R is the interval $\left[-\frac{\ln A(\tau)}{\tau} + \frac{B(\tau)}{\tau}r, \infty\right)$. Hence its density is zero outside this interval. For $x \in \left[-\frac{\ln A(\tau)}{\tau} + \frac{B(\tau)}{\tau}r, \infty\right)$ we have

$$\begin{aligned} F_R(x) &= \text{Prob}\left[-\frac{\ln A(\tau)}{\tau} + \frac{B(\tau)}{\tau}r + \frac{C(\tau)}{\tau}y < x\right] \\ &= \text{Prob}\left[y < \frac{\tau x + \ln A(\tau) - B(\tau)}{C(\tau)}\right] = F_y\left(\frac{\tau x + \ln A(\tau) - B(\tau)}{C(\tau)}\right) \end{aligned}$$

where we have used the positiveness of $C(\tau)$ again. Hence

$$f_R(x) = F_R'(x) = f_y\left(\frac{1}{C(\tau)}\left(\tau x + \ln A(\tau) - B(\tau)\right)\right)\frac{\tau}{C(\tau)},$$

$\diamondsuit$

Similarly as in the case of previously discussed two-factor models, also in the Fong-Vasicek model, the variances of bond prices and interest rates tend to zero as $\tau \to \infty$. However, in this case, we will need averaged values to prove it. For this reason, we postpone it to the next section.

9.3 Averaged bond prices and term structures. Confidence intervals.

Let us define the averaged bond prices and interest rates as

$$\begin{aligned} \tilde{P}(\tau, r) &= \langle P(\tau, r, y)\rangle_y, \\ \tilde{R}(\tau, r) &= \langle R(\tau, r, y)\rangle_y, \end{aligned}$$

where the expectations are taken with respect to the limiting distribution (9.9) of the process y.

Theorem 21. *[42] Averaged bond prices and interest rates, with respect to the limiting distribution of the random variable y are given by:*

1. $\tilde{P}(\tau, r) = A(\tau) e^{-B(\tau) r} \left(1 + \frac{C(\tau)}{\beta}\right)^{-\alpha}$,

2. $\tilde{R}(\tau, r) = -\frac{1}{\tau} \ln A(\tau) + \frac{B(\tau)}{\tau} r + \frac{C(\tau)}{\tau} \theta_2$.

Proof:

1. We compute the averaged bond price:

$$\begin{aligned}
\tilde{P}(\tau, r) &= \int_0^\infty P(\tau, r, y) g(y) dy \\
&= A(\tau) e^{-B(\tau) r} \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-(C(\tau)+\beta) y} y^{\alpha-1} dy \\
&= A(\tau) e^{-B(\tau) r} \frac{\beta^\alpha}{(C(\tau)+\beta)^\alpha} \\
&= A(\tau) e^{-B(\tau) r} \left(1 + \frac{C(\tau)}{\beta}\right)^{-\alpha}.
\end{aligned}$$

2. The formula for the averaged interest rate follows from linearity of interest rate in the y variable and taking into account the expected value of the limiting gamma distribution $\Gamma(\beta, \alpha) = \Gamma\left(\frac{2\kappa_2}{v^2}, \frac{2\kappa_2}{v^2}\theta_2\right)$ which is equal to θ_2.

$\diamondsuit$

Remark 7. *The function $y \to P(\tau, r, y)$ is strictly convex because its second derivative ∂_y^2 is equal to $C^2(\tau) P(\tau, r, y) > 0$. Hence, by Jensen's inequality, we have*

$$\tilde{P}(\tau, r) = \int_0^\infty P(\tau, r, y) g(y) dy > P\left(\tau, r, \int_0^\infty y g(y) dy\right) = P(\tau, r, \theta_2),$$

since $\int_0^\infty y g(y) dy = \theta_2$ i.e. the averaged bond price $\tilde{P}(\tau, r) = \langle P(\tau, r, y) \rangle_y$ is always greater than the bond price corresponding to the limiting mean value $\langle y \rangle_y = \theta_2$ of the stochastic volatility y.

Now, we are able to prove the assertion on limiting behavior of variances.

Theorem 22. *[42] For fixed values of τ and r we have*

$$\lim_{\tau\to\infty} Var_y P(\tau, r, y) = 0, \ \lim_{\tau\to\infty} Var_y R(\tau, r, y) = 0,$$

where variances are computed with respect to the limiting distribution of y.

Proof: We have already computed $\langle P(\tau,r,y)\rangle_y$. In the same way we compute the expected value of P^2:

$$\begin{aligned}\langle P^2(\tau,r,y)\rangle_y &= \int_0^\infty \left(A(\tau)e^{-B(\tau)r-C(\tau)y}\right)^2 g(y)dy\\ &= A(\tau)^2e^{-2B(\tau)r}\int_0^\infty e^{-(2C(\tau)+\beta)}y^{\alpha-1}dy\\ &= A(\tau)^2e^{-2B(\tau)r}\left(1+\frac{2C(\tau)}{\beta}\right)^{-\alpha}.\end{aligned}$$

Hence

$$\begin{aligned}Var_y(P(\tau,r,y)) &= \langle P^2(\tau,r,y)\rangle_y-\langle P(\tau,r,y)\rangle_y^2\\ &= A(\tau)^2e^{-2B(\tau)r}\left[\left(1+\frac{2C(\tau)}{\beta}\right)^{-\alpha}-\left(1+\frac{C(\tau)}{\beta}\right)^{-2\alpha}\right]\\ &= A(\tau)^2e^{-2B(\tau)r}\left[\left(1+\frac{2C(\tau)}{\beta}\right)^{-\alpha}-\left(1+\frac{2C(\tau)}{\beta}+\frac{C(\tau)^2}{\beta^2}\right)^{-\alpha}\right].\end{aligned}$$

By the mean value theorem

$$\left(1+\frac{2C(\tau)}{\beta}\right)^{-\alpha}-\left(1+\frac{2C(\tau)}{\beta}+\frac{C(\tau)^2}{\beta^2}\right)^{-\alpha}=\left(-\alpha\xi^{-\alpha-1}\right)\left(-\frac{C(\tau^2)}{\beta^2}\right)$$

for some ξ from the interval $\left(1+\frac{2C(\tau)}{\beta},1+\frac{2C(\tau)}{\beta}+\frac{C(\tau)^2}{\beta^2}\right)$. Hence $\xi>1$ and therefore

$$\begin{aligned}Var_y(P(\tau,r,y)) &= A(\tau)^2e^{-2B(\tau)r}\alpha\xi^{-\alpha-1}\frac{C(\tau^2)}{\beta^2}\\ &< A(\tau)^2e^{-2B(\tau)r}\frac{\alpha}{\beta^2}C(\tau^2).\end{aligned}$$

Since $C(\tau)$ and $B(\tau)$ are bounded on $[0,\infty)$ and $A(\tau)\to 0$ as $\tau\to\infty$, we conclude that

$$Var_y(P(\tau,r,y))\to 0 \text{ as } \tau\to\infty.$$

Since R is linear in the y variable and the variance of y is $Var(y)=\frac{\alpha}{\lambda^2}$, we obtain

$$Var_y(R(\tau,r,y))=\left(\frac{C(\tau)}{\tau}\right)^2 Var(y)=\frac{v^2\theta_2}{2\kappa_2}\frac{C^2(\tau)}{\tau^2}.$$

Because $C(\tau)$ is bounded and $\frac{1}{\tau^2}\to 0$ for $\tau\to\infty$ we obtain

$$Var_y(R(\tau,r,y))\to 0 \text{ for } \tau\to\infty,$$

as claimed in the theorem. $\diamondsuit$

As we know, $P(\tau, r, y)$ and $R((\tau, r, y)$ are monotone functions of y. Hence the area containing a given percentile of bond prices and term structures can be bounded by $P(\tau, r, y_1)$ and $P(\tau, r, y_2)$, respectively $R(\tau, r, y_1)$ and $R(\tau, r, y_2)$, where $\int_{y_1}^{y_2} g(y)dy$ equals the given percentile.

In Figure 9.1 we see examples of bond prices and term structures (corresponding to the same short rate but different volatilities) together with their averaged values and 95% confidence intervals.

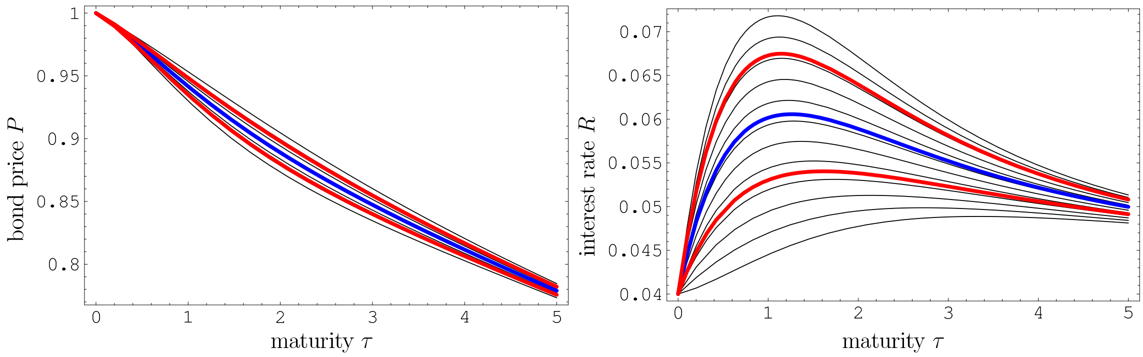

Figure 9.1: Bond prices (left) and term structures (right) in the Fong-Vasicek model (grey), their averaged values (blue) and confidence intervals (red).

9.4 Relation of averaged bond prices and one-factor models

We study a problem whether there are functions $\mu = \mu(r)$ and $\sigma = \sigma(r)$ such that the bond prices in the short rate model $dr = \mu dt + \sigma dw$ are the same as the averaged prices from the Fong-Vasicek model. We restrict ourselves to certain processes. Drift and volatility of the process, as well as the market price of risk λ are assumed to be time-independent. For zero level of the short rate, we require the volatility to be zero. This condition is needed for the nonnegativity of short rate.

Theorem 23. *[46, Theorem 1] Consider the averaged bond prices $\tilde{P}(\tau, r)$ in Fong-Vasicek model and the following conditions on the one-factor model:*

1. *functions μ, σ, λ depend only on the short rate r and not on time t,*
2. *functions μ, σ, λ are continuous in $r = 0$, $\sigma(0) = 0$.*

Then there is no such one-factor interest rate model, for which the averaged bond prices satisfy the bond pricing PDE

$$-\frac{\partial P}{\partial \tau} + (\mu(r) - \lambda(r)\sigma(r)\frac{\partial P}{\partial r} + \frac{1}{2}\sigma(r)^2\frac{\partial^2 P}{\partial r^2} - rP = 0 \qquad (9.12)$$

for $r \geq 0$ *and* $\tau > 0$.

Proof: Suppose that such a one-factor model exists. To be able to insert the averaged price $\tilde{P}(\tau, r)$ into the PDE (9.12) for bond prices, It follows from Theorem 21 that we start with computing the necessary partial derivatives of $\tilde{P}$:

$$\begin{aligned}
\frac{\partial \tilde{P}}{\partial \tau} &= \left(\frac{A'}{A} - B'r - \frac{\alpha C'}{\beta + C}\right)\tilde{P}, \\
\frac{\partial \tilde{P}}{\partial r} &= -B\tilde{P}, \\
\frac{\partial^2 \tilde{P}}{\partial r^2} &= B^2\tilde{P}.
\end{aligned}$$

We use a similar idea as in Theorem 17 for the case of two-factor CIR model. We suppose that the averaged bond price $\tilde{P}(\tau, r)$ satisfies the one-factor PDE for bond prices. Then, when taking a limit $r \to 0^+$, the terms involving volatility vanish, since $\sigma(0) = 0$. Terms involving drift converge to $\mu(0)$, which has to be constant with respect to τ. In the case of the Fong-Vasicek model we obtain

$$\mu(0) = \left.\frac{\partial_\tau \tilde{P}}{\partial_r \tilde{P}}\right|_{r=0} = \left.\frac{\frac{A'}{A} - B'r - \frac{\alpha}{\beta+C}C'}{-B}\right|_{r=0} = \frac{\frac{A'}{A} - \frac{\alpha}{\beta+C}C'}{-B}.$$

Hence the necessary condition for $\tilde{P}$ to be a bond price in a one-factor model is that:

$$\frac{-\kappa_1\theta_1 B - \kappa_2\theta_2 C - C'\frac{\alpha}{\beta+C}}{-B} = k$$

for all $\tau > 0$ and some constant $0 < k < \infty$. Then

$$-\kappa_1\theta_1 B - \kappa_2\theta_2 C - \alpha\frac{C'}{\beta + C} + kB = 0 \tag{9.13}$$

for all $\tau > 0$. Hence also the derivative with respect to τ of the left hand side is identically zero, so

$$-\kappa_1\theta_1 B' - \kappa_2\theta_2 C' - \alpha\frac{(\beta + C)C'' - (C')^2}{(\beta + C)^2} + kB' = 0.$$

Since this equality holds for all $\tau > 0$, also the limit of its left hand side for $\tau \to 0+$ equals zero. Using the initial condition for C and the values of its derivatives for $\tau = 0$ (see Theorem 19), it yields

$$\kappa_1\theta_1 - \lambda_1\theta_2 = k.$$

Substituting this expression for k into (9.13) gives

$$-\kappa_2\theta_2 C - \frac{\alpha}{\beta + C}C' - \lambda_2\theta_2 B = 0,$$

from which we express $\alpha C'$ as

$$\alpha C' = (\beta + C)(\kappa_2\theta_2 C - \lambda_1\theta_2 B). \tag{9.14}$$

On the other hand, from (9.6) we know that

$$\alpha C' = \alpha\left(-\lambda_1 B - \frac{B^2}{2} - (\kappa_2 + \lambda_2 v + v\rho B)\,C - \frac{v^2}{2}C^2\right). \tag{9.15}$$

Hence the right hand sides of (9.14) and (9.15) have to be equal. From this equality, using the relation between parameters α and β, we are able to express the function C explicitly as

$$C(\tau) = \frac{\alpha\left(\lambda_2 - \lambda_1 - \frac{\beta}{2}\right)}{v\alpha\rho - \lambda_2\theta_2 + \lambda_2 v\frac{1}{B(\tau)}},$$

where $B(\tau)$ is given by (9.5). The derivative $C'(\tau)$ can be computed to be

$$C'(\tau) = -\frac{2e^{\kappa_1\tau}\kappa_1^2\kappa_2\lambda_2\theta_2(\kappa_2 + (\lambda_1 - \lambda_2)v^2)}{v\left(2(e^{\kappa_1\tau} - 1)\kappa_2\rho\theta_2 + \lambda_2 v(\theta_2 - e^{\kappa_1\tau}\theta_2 + e^{\kappa_1\tau}\kappa_1 v\right)^2}.$$

and hence

$$C'(0) = \frac{2\kappa_2\theta_2(\kappa_2 + (\lambda_1 - \lambda_2)v^2)}{\lambda_2 v^5}.$$

We already know that $C'(0) = 0$. Since κ_2 and θ_2 are positive parameters, it implies that $\kappa_2 + (\lambda_1 - \lambda_2)v^2 = 0$. But then $C'(\tau)$ would be identically zero and hence $C(\tau)$ would be constant with respect to τ, which is a contradiction. $\diamondsuit$

9.5 Fast mean reverting volatility

It was observed on the several financial markets (interest rates, stocks, stock indices) that the stochastic volatility evolves in a different time scale, which is faster than the scale of the underlying security. For a concise reference of the asymptotic we refer the reader to the book [25]. In this section we are interested in fast time scale of volatility in the Fong-Vasicek model.

If the volatility evolves in a time scale with the unit $\varepsilon > 0$, then the stochastic differential equations (9.1) for r and y in become

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}dw_1, \\ dy &= \frac{\tilde{\kappa_2}}{\varepsilon}(\theta_2 - y)dt + \frac{\tilde{v}}{\sqrt{\varepsilon}}\sqrt{y}dw_2. \end{aligned}$$

Scaling of the volatility $\tilde{v}$ by the factor $\sqrt{\varepsilon}$ is because the order or the stochastic term: we have $dw = \Phi\sqrt{dt}$, where $\Phi \sim N(0,1)$. The fast mean reverting volatility corresponds to small values of ε and the limit $\varepsilon \to 0$. If we define $\kappa_2 = \tilde{\kappa_2}/\varepsilon$, $v = \tilde{v}/\sqrt{\varepsilon}$, we obtain the stochastic differential equation

$$dy = \kappa_2(\theta_2 - y)dt + v\sqrt{y}dw_2,$$

where the ratio κ_2/v^2 is constant with respect to ε. Hence we can fix the ratio

$$\tilde{\kappa_2}/\tilde{v}^2 = \kappa_2/v^2 = k \tag{9.16}$$

and consider the fast mean reversion as a limit $v \to \infty$.

First, we present a numerical example. In Figure 9.2, there are terms structures for the same values of parameters κ_1, θ_1, θ_2 and k, the same values of short rate and volatility, but the increasing speed of volatility evolution. We notice that the differences between the interest rates decrease with increasing speed of volatility. In the rest of this section, we prove this observation analytically.

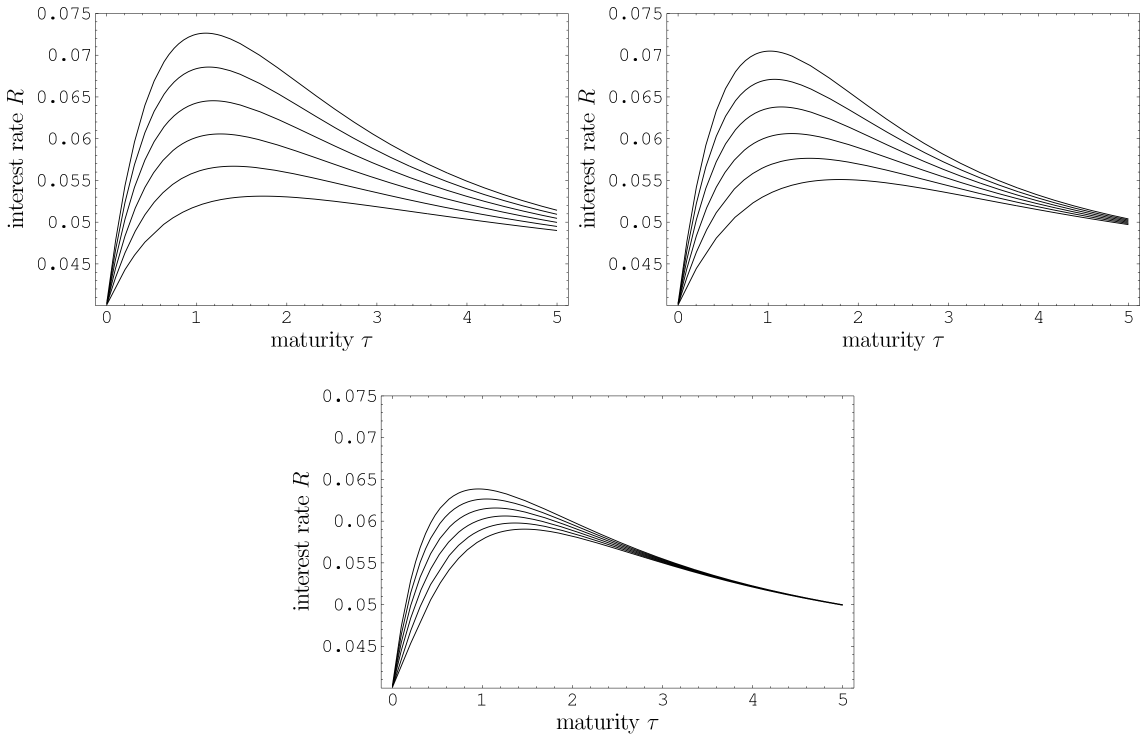

Figure 9.2: Term structures for the same values of y and increasing speed of volatility. Term structures in each graph correspond to the same volatility levels y. However, the graphs have different time scales, the speed increases from left to right.

Recall that the price P of a bond can be written in a form

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

where A, B, C are functions of τ and satisfying (9.4). Since we work with different values of v now, functions A, C will be functions of τ, as well as of the parameter v: $A = A(\tau, v)$, $C = C(\tau, v)$. Function $B = B(\tau)$ does not depend on value of v.

Let us define a function

$$D(\tau, v) = \frac{\partial C}{\partial v}(\tau, v). \tag{9.17}$$

Substituting (9.16) into the equation (9.6) for C we obtain

$$\begin{cases} \frac{\partial C}{\partial \tau} + \lambda_1 B + \frac{1}{2}B^2 + (kv^2 C + \lambda_2 v + v\rho B)C + \frac{v^2 C(\tau)^2}{2} = 0, \\ C(0, v) = 0, \end{cases}$$

from which by taking derivative with respect to v we obtain equation for $D(\tau, v)$:

$$\begin{cases} \frac{\partial D}{\partial \tau} + (2kv + \lambda_2 + \rho B)C + (kv^2 C + \lambda_2 v + v\rho B)D + vC^2 + v^2 CD = 0, \\ D(0, v) = 0. \end{cases}$$

Using this equation, which is satisfied by D, we prove the following result.

Theorem 24. *[44] There exists $v_0 > 0$ such that $D(\tau, v) < 0$ for all $v > v_0$ and $\tau > 0$.*

Proof: First we prove a usedull property of the function D: If $D(\tau, v) = 0$, then $\frac{\partial d}{\partial \tau} D(\tau, v) < 0$. Indeed, if $D(\tau, v) = 0$ for some v and $\tau > 0$, then

$$\frac{\partial D}{\partial \tau} = -\left[(2kv + \lambda_2 + \rho B)C + vC^2\right].$$

Since the function B is bounded with a bound independent on v, we can find $v_0 > 0$ such that for all $v > v_0$ we have

$$2kv + \lambda_2 + \rho B(\tau) > 0$$

for any $\tau \geq 0$. Hence, if $D(\tau, v) = 0$ for some $v > v_0$ then

$$\frac{\partial D}{\partial \tau}(\tau, v) < 0.$$

Let us fix $v > v_0$ and consider D as a function of the variable τ. Taking into account the above property of D, to prove the claim of the theorem, it suffices to show that $D(\tau, v) < 0$ for τ on some neighborhood of $\tau = 0$. From the previous parts we know that $C(0) = 0,\ C'(0) = 0,\ C''(0) = -\lambda_1 > 0$. The function $D(\tau, v)$ satisfies $D(0, v) = 0$. Taking the time derivatives from the equation for D gives

$$D'(0, v) = 0,\ D''(0, v) = 0,\ D'''(0, v) = -(2kv + \lambda_2 + \rho)C''(0) < 0.$$

Hence we have

$$D(0, v) = 0,\ D'(0, v) = 0,\ D''(0, v) = 0,\ D'''(0, v) < 0,$$

which implies that $D(\tau, v) < 0$ on a neighborhood of $\tau = 0$. $\diamondsuit$

Lemma 3. *[44] Let $\tau > 0$. Then:*

1. *There exists $v_0 > 0$ such that for $v > v_0$, the function $C(\tau, v)$ is decreasing in variable v.*

2. *There exists a finite limit $L(\tau) = \lim_{v\to\infty} C(\tau, v)$. Then $L(\tau) \geq 0$ for all $\tau > 0$.*

Proof:

1. Follows from Theorem 24, as $\frac{\partial C}{\partial v} = D < 0$

2. For $v > v_0$ the function $C(\tau, v)$ is decreasing and it is bounded by zero from below. Hence the limit $\lim_{v\to\infty} C(\tau, v)$ exists and it is nonnegative.

$\diamondsuit$

In what follows, we will show that if $L(\tilde{\tau}) > 0$ for some $\tilde{\tau} > 0$, then it would lead to an unrealistic behavior of interest rates. Firstly we show that in this case $L(\tau) > 0$ on some interval $(\tilde{\tau} - h, \tilde{\tau})$. We have

$$\frac{\partial C}{\partial \tau}(\tilde{\tau}, v) = -\left[(\lambda_1 B + \frac{1}{2}B^2) + v(\lambda_2 + \rho B)C + v^2(k + \frac{1}{2}C)C\right] \to -\infty$$

for $v \to \infty$, which implies that

$$\frac{\partial C}{\partial \tau}(\tilde{\tau}, v) < 0,$$

for $v > v_0$, where v_0 is sufficiently large and hence

$$C(\tau, v) > C(\tilde{\tau}, v) \text{ for } \tau \in (\tilde{\tau} - h, \tilde{\tau}),$$

where $h > 0$ is sufficiently small. Taking the limit $v \to \infty$ we obtain

$$L(\tau) \geq L(\tilde{\tau}),$$

and for $\tau \in (\tilde{\tau} - h, \tilde{\tau})$ we have $L(\tau) > 0$ as well.

Now consider τ such that $(0, \tau)$ contains the interval $(\tilde{\tau} - h, \tilde{\tau})$. For interest rates with this maturity we then have:

$$\begin{aligned} R(\tau, r, y, v) &= \frac{\log A(\tau)}{\tau} + \frac{B(\tau)}{\tau}r + \frac{C(\tau, v)}{\tau}y \\ &= \theta_1 - \theta_1\frac{B(\tau)}{\tau} + \frac{1}{\tau}\theta_2 k v^2 \int_0^\tau C(s, v)ds + \frac{B(\tau)}{\tau}r + \frac{C(\tau, v)}{\tau}y. \end{aligned}$$

The following equality holds:

$$\lim_{v\to\infty} \int_0^\tau C(s, v)ds = \int_0^\tau \lim_{v\to\infty} C(s, v) = \int_0^\tau L(s)ds > 0,$$

since changing order of limit and integration is justified by monotony of the function $C(\tau, v)$ in the variable v for $v > v_0$. It means that

$$R(\tau, r, y, v) \to \infty \text{ for } v \to \infty.$$

However, although the fast mean reversion for market data is empirically confirmed (c.f. [25]), the limit above is not an observed property.

We now analyze the second case left when $L(\tau) = 0$ for all $\tau > 0$. Let us consider two values of the volatility y, without loss of generality we can assume that $y_1 < y_2$. Then $R(\tau, r, y_1, v) < R(\tau, r, y_2, v)$ and the difference between the two interest rates is

$$R(\tau, r, y_2, v) - R(\tau, r, y_1, v) = C(\tau, v)\frac{y_2 - y_1}{\tau}.$$

As a function of the variable v, the term $(y_2 - y_1)/\tau$ is a positive constant, and $C(\tau, v)$ is a decreasing function for large values of v. Hence also the difference between the interest rates decreases with increasing value of the parameter v. Recalling that the limit for $v \to \infty$ is $L(\tau) = 0$, we conclude that the differences between the interest rates converge to zero and this convergence is monotone for large values of v.

9.6 Volatility clustering

As we have already seen, in the Fong-Vasicek model the volatility has a limit distribution with a density having one maximum. It was a consequence of properties of mean-reversion process for its evolution. Now, we are looking for a model in which the limiting density has two local maxima. This corresponds to the so called volatility clustering, when the volatility can be in its high level (taking values around one of the local maxima of the limiting distribution) and its low level (taking values around the other local maximum). The desired behavior of the process and its limiting density are show in the Figure 9.3.

A natural candidate for such a volatility process y is

$$dy = a(y)dt + \sigma(y)dw,$$

which has a drift $a(y)$ such that the differential equation $dy = a(y)dt$ has two stable stationary solutions. With added stochastic part of the process, these stationary solutions become values, around which the volatility concentrates.

We propose a model with this property, for which the limiting density is a combination of two gamma densities. Consider the following two stochastic processes:

$$\begin{aligned} dy_1 &= a_1(y_1)dt + \sigma\sqrt{y_1}dw_1, \\ dy_2 &= a_2(y_2)dt + \sigma\sqrt{y_2}dw_2, \end{aligned}$$

where

$$a_1(y) = \kappa(\theta_1 - y), \; a_2(y) = \kappa(\theta_2 - y),$$

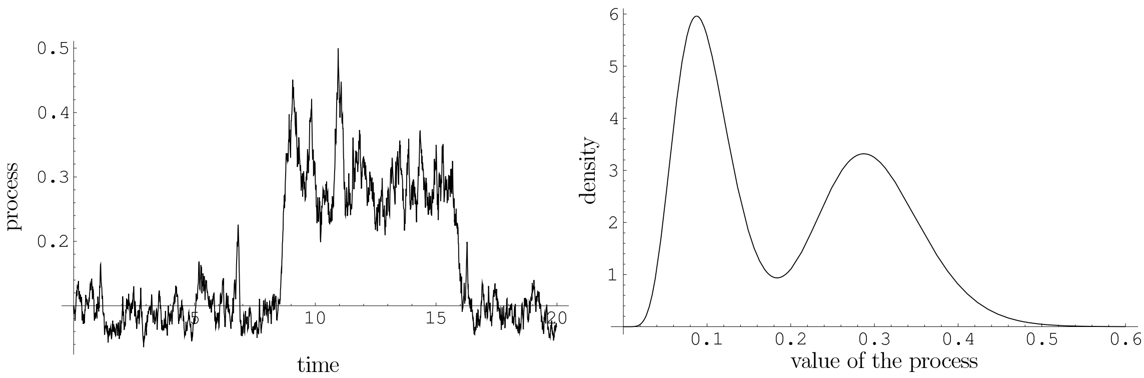

Figure 9.3: Simulation of a process in the left and its asymptotic distribution in the right.

such that $\theta_1 < \theta_2$, $2\kappa\theta_1 \leq \sigma^2$, $2\kappa\theta_2 \leq \sigma^2$. Their limit distributions are gamma distributions $\Gamma(\alpha, \alpha\theta_1)$, resp. $\Gamma(\alpha, \alpha\theta_2)$, where $\alpha = 2\kappa/\sigma^2$. Denote their densities by g_1 andg_2. Choose $k \in (0,1)$; our aim is to construct a process with asymptotic density

$$g(y) = kg_1(y) + (1-k)g_2(y), \tag{9.18}$$

corresponding to a mixture of densities g_1 and g_2.

In the following theorem we show that for the same volatility function (i.e. $\sigma\sqrt{y}$) it is possible to achieve this goal. Drift of the process can be written as a weighted sum of a_1 a a_2, with the weights dependent on y. We find the behavior of this function for $y \to 0$ and $y \to \infty$.

Theorem 25. *[43]*

1. *A process $dy = a(y)dt + \sigma\sqrt{y}dw$ has a limit distribution (9.18) for*

$$a(y) = w(y)a_1(y) + (1 - w(y))a_2(y),$$

where

$$w(y) = \frac{kg_1(y)}{kg_1(y) + (1-k)g_2(y)}.$$

2. *The function a satisfies:*

$$\lim_{y\to 0} \frac{a_1(y)}{a(y)} = 1, \ \lim_{y\to\infty} \frac{a_2(y)}{a(y)} = 1.$$

Proof:

1. We recall (c.f. for example [27]) that if the process $y(t)$ satisfies a stochastic differential equation

$$dy(t) = a(t, y(t))dt + b(t, y(t))dw, \tag{9.19}$$

then the conditional density $f(t, y|y(0) = y_0)$ of the random variable $y(t)$ satisfying an initial condition $y(0) = y_0$ is a solution to the Fokker-Planck partial differential equation

$$-\frac{\partial f}{\partial t} + \frac{\partial(af)}{\partial y} + \frac{\partial^2}{\partial y^2}\left(\frac{b^2 f}{2}\right) = 0 \tag{9.20}$$

for $t > 0$ with the initial condition $f(0, y) = \delta_0(y - y_0)$ where δ_0 is the Dirac function. The limiting distribution $g(y) = \lim_{t\to\infty} f(t, y)$ can be computed by solving the stationary Fokker-Planck equation.

$$-\frac{d}{dy}(ag) + \frac{d^2}{dy^2}\left(\frac{b^2 f}{2}\right) = 0 \tag{9.21}$$

with the normalization condition

$$\int_{-\infty}^{\infty} g(y)dy = 1.$$

Now, consider the positive process with the volatility $b(y) = v\sqrt{y}$, such that $\lim_{y\to 0+} g(y) = 0$ and $g'(y)$ is bounded on some neighborhood of the origin. Then its asymptotic distribution $g(y)$ is a solution of

$$-a(y)g(y) + \frac{d}{dy}\left(\frac{v^2 y g(y)}{2}\right) = 0.$$

From this, we compute the drift $a(y)$:

$$\begin{aligned} a(y) &= \frac{1}{g(y)}\left(\frac{v^2 y g(y)}{2}\right) = \frac{1}{g(y)}\frac{d}{dy}\left(\frac{v^2 y g(y)}{2}\left(kg_1(y) + (1-k)g_2(y)\right)\right) \\ &= \frac{1}{g_1(y)}\left(k\frac{d}{dy}\left(\frac{v^2 y g(y)}{2}\right) + (1-k)\frac{d}{dy}\left(\frac{v^2 y g_2(y)}{2}\right)\right). \end{aligned}$$

Since

$$\frac{d}{dy}\left(\frac{v^2 y g_1(y)}{2}\right) = a_1(y)g_1(y), \quad \frac{d}{dy}\left(\frac{v^2 y g_2(y)}{2}\right) = a_2(y)g_2(y),$$

$a(y)$ can be written as

$$a(y) = w(y)a_1(y) + (1 - w(y))a_2(y), \tag{9.22}$$

where

$$w(y) = \frac{kg_1(y)}{kg_1(y) + (1-k)g_2(y)}. \tag{9.23}$$

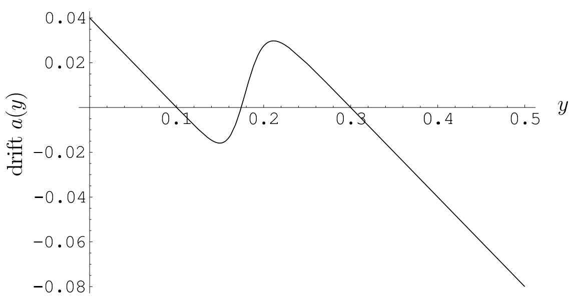

Figure 9.4: The resulting drift function $a(y)$.

2. Substituting g_1 a g_2 into (9.23) we obtain the weight $w(y)$:

$$w(y) = \frac{1}{1 + \frac{1-k}{k}\frac{g_2(y)}{g_1(y)}} = \frac{1}{1 + \frac{1-k}{k}\frac{\frac{\alpha^{\alpha\theta_2}}{\Gamma(\alpha\theta_2)}y^{\alpha\theta_2 - 1}e^{-\alpha y}}{\frac{\alpha^{\alpha\theta_1}}{\Gamma(\alpha\theta_1)}y^{\alpha\theta_1 - 1}e^{-\alpha y}}} = \frac{1}{1 + cy^q}, \tag{9.24}$$

where

$$c = \frac{1-k}{k}\frac{\Gamma(\alpha\theta_1)}{\Gamma(\alpha\theta_2)}\alpha^{\alpha(\theta_2-\theta_1)} > 0,\ q = \alpha(\theta_2 - \theta_1) > 0. \tag{9.25}$$

It follows that $w(y) \to 1$ for $y \to 0$ and $w(y) \to 0$ for $y \to \infty$. Therefore $\lim_{y\to 0}\frac{a_1(y)}{a(y)} = 1$ and similarly $\lim_{y\to 0}\frac{a_2(y)}{a(y)} = 1$, as claimed. $\diamondsuit$

In Figure 9.4 we can see an example of a drift function $a(y)$ obtained by Theorem 25.

As we have already mentioned at the beginning of the section, the expected behavior of function $a(y)$ is such, that it generates two stable fixed points of the ODE $dy = a(y)dt$. Function depicted in Figure 9.4 has this property. There are three points y for which $a(y) = 0$. Two of them with negative slope of $a(y)$ correspond to desired stable fixed points and there is one unstable among them. In Figure 9.5 we see that this is not a property, which necessarily holds for a general case of the function $a(y)$. However, we show that for large values of parameter α (i.e. sufficiently strong mean reversion of the processes which we construct the process from), the function $a(y)$ has the expected behavior.

Theorem 26. *[43] There exists α_0 such that for $\alpha > \alpha_0$ the equality $a(y) = 0$ for exactly three values of y.*

Proof: According to Theorem 25 function $a(y)$ can be written as

$$a(y) = \frac{\kappa}{1 + cy^q}\left[(\theta_1 - y) + cy^q(\theta_2) - y)\right]. \tag{9.26}$$

We see that $a(y) > 0$ for $y \in [0, \theta_1]$ and $a(y) < 0$ for $y \in [\theta_2, \infty)$. It follows that all the points for which the function a is zero, are from the interval (θ_1, θ_2).

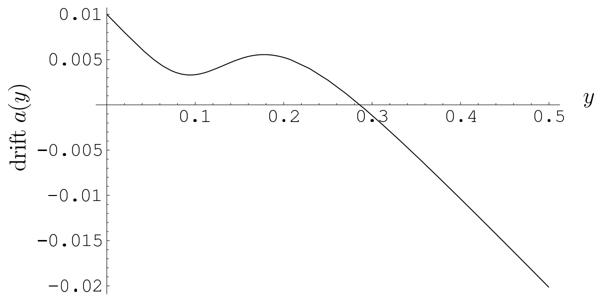

Figure 9.5: Example of a function $a(y)$, which does not have three zero points and, consequently, the ODE $dy = a(y)dt$ does not have the expected equilibria.

Firstly, we show that there are at most three such points. From (9.26) we see that $a(y) = 0$ if and only if the function

$$f(y) = (\theta_1 - y) + cy^q(\theta_2 - y). \tag{9.27}$$

equals zero. Suppose that $f(y) = 0$ in k points. Using Rolle's theorem we obtain that there are at least $k-2$ points in interval (θ_1, θ_2), for which $f''(y) = 0$. But

$$f''(y) = cq(q-1)\theta_2 y^{q-2} - c(q+1)qy^{q-1},$$

and therefore there exists at most one such point $y = \theta_2(q-1)/(q+1)$ or none, if this value is less than θ_1. Hence $k \leq 3$.

Now, we rewrite the equality $f(y) = 0$ as

$$cy^q = \frac{y - \theta_1}{\theta_2 - y}. \tag{9.28}$$

By defining the functions

$$f_1(y) = cy^q, \; f_2(y) = \frac{y - \theta_1}{\theta_2 - y}, \quad y \in (\theta_1, \theta_2)$$

it suffices to find $y_1 < y_2$ such that

$$f_1(y_1) < f_2(y_1), \; f_1(y_2) > f_2(y_2). \tag{9.29}$$

Then (9.28) holds for exactly three points y.

We find the limit of the function f_1 for fixed y as $\alpha \to \infty$. For $x > 1$ we have the Stirling's formula (c.f. [1, formula (6.1.38)])

$$\Gamma(x) = \sqrt{2\pi}(x-1)^{x-1/2}e^{-(x-1)+\frac{1}{12(x-1)}\xi}$$

for some $\xi \in (0,1)$. Since we are interested in limit for $\alpha \to \infty$, we can assume that $\alpha\theta_1 > 1, \alpha\theta_2 > 1$. Then

$$\begin{aligned} f_1(y) &= cy^q = \frac{1-k}{k}\frac{\Gamma(\alpha\theta_1)}{\Gamma(\alpha\theta_2)}(\alpha y)^{\alpha(\theta_2-\theta_1)} = \\ &= \frac{1-k}{k}\frac{\sqrt{2\pi}(\alpha\theta_1-1)^{\alpha\theta_1-1/2}e^{-(\alpha\theta_1-1)+\frac{1}{12(\alpha\theta_1-1)}\xi_1(\alpha)}}{\sqrt{2\pi}(\alpha\theta_2-1)^{\alpha\theta_2-1/2}e^{-(\alpha\theta_2-1)+\frac{1}{12(\alpha\theta_2+1)}\xi_2(\alpha)}}\frac{(\alpha y)^{\alpha\theta_2}}{(\alpha y)^{\alpha\theta_1}}, \end{aligned}$$

where ξ_1, ξ_2 (dependent on α), are in the interval $(0,1)$. This can be written as

$$f_1(y) = \left[\frac{1-k}{k}\left(\frac{\alpha\theta_1-1}{\alpha\theta_2-1}\right)^{-\frac{1}{2}}\frac{\left(1-\frac{a}{\alpha\theta_1}\right)^{\alpha\theta_1}}{\left(1-\frac{a}{\alpha\theta_2}\right)^{\alpha\theta_2}}\frac{e^{\frac{\xi_1(\alpha)}{12(\alpha\theta_1-1)}}}{e^{\frac{\xi_2(\alpha)}{12(\alpha\theta_2-1)}}}\right]\left(\frac{\theta_1}{y}\right)^{\alpha\theta_1}\left(\frac{y}{\theta_2}\right)^{\alpha\theta_2}e^{\alpha(\theta_2-\theta_1)}.$$

Since

$$\lim_{\alpha\to\infty}\left[\frac{1-k}{k}\left(\frac{\alpha\theta_1-1}{\alpha\theta_2-1}\right)^{-\frac{1}{2}}\frac{\left(1-\frac{a}{\alpha\theta_1}\right)^{\alpha\theta_1}}{\left(1-\frac{a}{\alpha\theta_2}\right)^{\alpha\theta_2}}\frac{e^{\frac{\xi_1(\alpha)}{12(\alpha\theta_1-1)}}}{e^{\frac{\xi_2(\alpha)}{12(\alpha\theta_2-1)}}}\right] = \frac{1-k}{k}\left(\frac{\theta_1}{\theta_2}\right)^{-1/2} > 0,$$

we need to compute the limit

$$\lim_{\alpha\to\infty}\left(\frac{\theta_1}{y}\right)^{\alpha\theta_1}\left(\frac{y}{\theta_2}\right)^{\alpha\theta_2}e^{\alpha(\theta_2-\theta_1)} = \lim_{\alpha\to\infty}\left(e^{\theta_1\left(\log\frac{\theta_1}{y}\right)+\theta_2\left(\log\frac{y}{\theta_2}\right)+\theta_2-\theta_1}\right)^{\alpha}.$$

Denote

$$h(y) = \theta_1\left(\log\frac{\theta_1}{y}\right) + \theta_2\left(\log\frac{y}{\theta_2}\right) + \theta_2 - \theta_1.$$

For $x > -1, x \neq 0$ we have $\log(1+x) < x$. Hence, from the continuity of h it follows that there exist $\tilde{\theta}_1, \tilde{\theta}_2$, such that $\tilde{\theta}_1 < \tilde{\theta}_2$ and $h(y) < 0$ for $y \in (\theta_1, \tilde{\theta}_1)$, $h(y) > 0$ for $y \in (\tilde{\theta}_2, \theta_2)$. Then:

$$\begin{aligned} h(\theta_1) &= \theta_2\left(1-\frac{\theta_1}{\theta_2}+\log\frac{\theta_1}{\theta_2}\right) < 0, \\ h(\theta_2) &= -\theta_1\left(1-\frac{\theta_2}{\theta_1}+\log\frac{\theta_2}{\theta_1}\right) > 0. \end{aligned}$$

From this we obtain:

- For $y \in (\theta_1, \tilde{\theta}_1)$:

$$\lim_{\alpha\to\infty}\left(e^{\theta_1\left(\log\frac{\theta_1}{y}\right)+\theta_2\left(\log\frac{y}{\theta_2}\right)+\theta_2-\theta_1}\right)^{\alpha} = \lim_{\alpha\to\infty}\left(e^{h(y)}\right)^{\alpha} = 0,$$

  and hence for $f_1 = f_1(y,\alpha)$

$$\lim_{\alpha\to\infty} f_1(y,\alpha) = 0.$$

- For $y \in (\tilde{\theta}_2, \theta_2)$:

$$\lim_{\alpha\to\infty}\left(e^{\theta_1\left(\log\frac{\theta_1}{y}\right)+\theta_2\left(\log\frac{y}{\theta_2}\right)+\theta_2-\theta_1}\right)^{\alpha} = \lim_{\alpha\to\infty}\left(e^{h(y)}\right)^{\alpha} = \infty,$$

  and hence

$$\lim_{\alpha\to\infty} f_1(y,\alpha) = \infty.$$

Choose $y_1 \in (\theta_1, \tilde{\theta}_1)$ and $y_2 \in (\tilde{\theta}_2, \theta_2)$. Since the function f_2 does not depend on α, from the limits that we have computed it follows that there are α_1, α_2 such that

$$f_1(y_1) < f_2(y_1) \text{ for } \alpha > \alpha_1,$$

$$f_1(y_2) > f_2(y_2) \text{ for } \alpha > \alpha_2.$$

Hence for $\alpha > \max\{\alpha_1, \alpha_2\}$ the inequalities (9.29) are satisfied, from which the theorem follows. $\diamondsuit$

To provide an empirical evidence for such a volatility process, we computed maximum likelihood estimates of the volatility σ for the Vasicek model for each month from the period 2000-2007 using Bribor overnight data. Figure 9.6 shows the estimated as a function of time. The higher and lower volatility periods can be distinguished. They can be seen also on the histogram and kernel density estimates of the values in Figure 9.7.

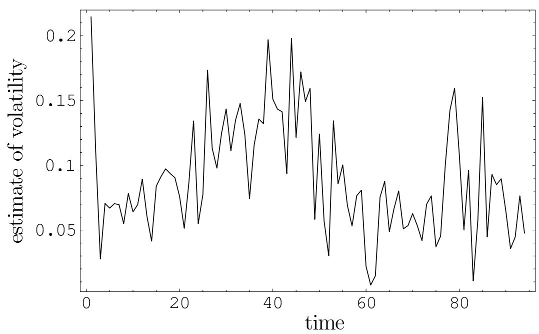

Figure 9.6: Monthly estimates of Vasicek model's σ (Bribor, 2000-2007).

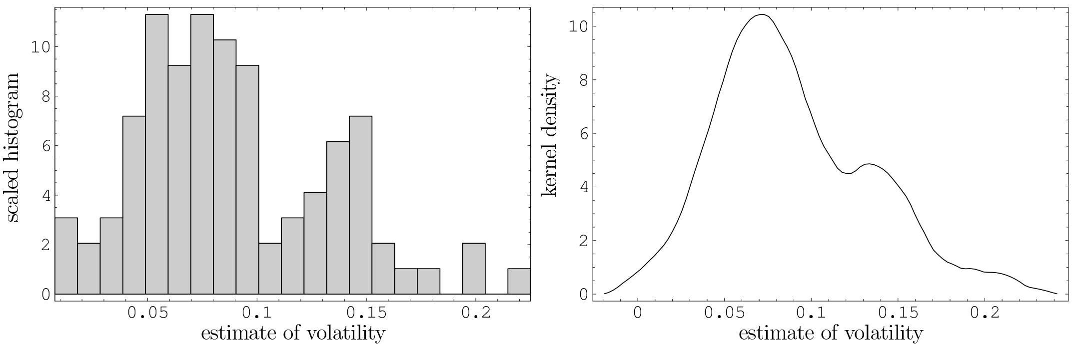

Figure 9.7: The histogram and kernel density of estimates of the volatility σ.

Chapter 10

Conclusion

We studied several questions and problems that are related to short rate term structure models. The main results of this thesis can be summarized as follows:

First we considered an approximate analytical solution for bond prices in one-factor models, suggested by Choi and Wirjanto in [17]. The presented approximation results are based on the recent paper [47]. We derived order of its accuracy and proposed an approximation of higher order. The approximation was used in calibration of the models and finding an optimal dependence of the volatility on the short rate.

Next we studied two-factor models. Namely, two-factor Vasicek, two-factor Cox-Ingersoll-Ross and Fong-Vasicek models in particular. We considered the limit distribution of the unobservable variables of the model. With respect to this distribution we studied the bond prices and interest rates. Most of the results are contained in the author's papers [42], [45] and [46]. They include deriving distribution of bond prices and interest rates, dependence of their variance on maturity, computing their averaged values and confidence intervals. We also studied the question of the nonexistence of one-factor models, yielding the same bond prices as are the averaged values from these models.

In the case of Vasicek model we furthemore studied the fast mean reverting volatility, based on author's paper [44]. We analyzed a condition enabling us to eliminate a possibility of infinite limit of interest rates for fast time scale of volatility. In this case, we have shown the monotone decreasing dependence of the difference between the interest rates, as the speed of volatility increases.

In the last section we presented the results of the author's paper [43]. We derived a process modeling volatility clustering. Its limit distribution is a mixture of two gamma densities. It generalizes gamma distribution from the Fong-Vasicek model.

Chapter 11

List of symbols

a, b - parameters of a drift function (chapter 4)
a, b - parameters of gamma distribution (chapter 8)
a, a_1, a_2 - drift functions (chapter 9)
A, B, A_1, A_2, B_1, B_2, C - shape functions of the interest rates
$Corr(.,.)$ - correlation of random variables
$E(.)$ - expected value of a random variable
EOC - experimental order of convergence
$f(.)$, $g(.)$ - density functions
F, $F(\alpha)$, $F(\alpha, \beta)$ - objective function in the calibration (chapter 5)
$F(.)$ - cumulative distribution function
$N(.,.)$ - normal distribution with specified parameters
P, $P(t, T, r)$, $P(\tau, r)$, $P(\tau, r_1, r_2)$, $P(\tau, r, y)$ - price of bond
P^{ap}, P^{ap2} - approximations of bond prices
$Prob(.)$ - probability of an event
$\tilde{P}$, $\tilde{P}(\tau, r)$ - averaged bond price
r - short rate
r_1, r_2 - factors of the short rate
R, $R(t, T, r)$, $R(\tau, r)$, $R(\tau, r_1, r_2)$, $R(\tau, r, y)$ - interest rate
R^{ap}, R^{ap2} - approximations of interest rates
$\tilde{R}$, $\tilde{R}(\tau, r)$ - averaged interest rate
T - maturity of a bond
$Var(.)$ - variance of a random variable
$Var(.|.)$ - conditional variance of a random variable

w, $w(t)$, w_t - Wiener process
w_{ij} - weights in the calibration (chapter 5)
$w(.)$ - weight function for drifts in volatility clustering model (chapter 9)
α, β - parameters of a drift function (chapter 4)
α, β - parameters of gamma distribution (chapter 9)
$\gamma > 0$ - parameter describing the dependence of volatility on the short rate
$\Gamma(.)$ - gamma function
$\Gamma(.,.)$ - gamma distribution with specified parameters
$\theta > 0$ - long time limit of a mean reversion a process
$\kappa > 0$ - mean reversion parameter of a process
λ, $\lambda(t,r)$ $\lambda(r)$, $\tilde{\lambda}$ - market price of risk
μ, $\mu(t,r)$, $\mu(r)$ - drift function of a process
μ, σ^2 - parameters of normal distribution (chapter 7)
σ, $\sigma(t,r)$, $\sigma(r)$ - volatility function of a process
$\sigma > 0$, $v > 0$ - volatility parameter of a process
τ - time to maturity
${}_1F_1(.,.,.)$ - Kummer confluent hypergeometric function
$\langle . \rangle$ - expected value of a random variable
$\langle . | . \rangle$ - conditional expected value of a random variable

Bibliography

[1] Abramovitz, M., Stegun. I.: Handbook on Mathematical functions. National Bureau of Standards, Applied Mathematics series 55, 10th edition, 1972.

[2] Anděl, J.: Základy matematické statistiky. Matfyzpres, Praha, 2005.

[3] Adkins, L. C., Krenbiel, T.: Mean reversion and volatility of short-term London Interbank Offered Rates: An empirical comparison of competing models. International Review of Economics and Finance 8, 1997, 45-55.

[4] Anderson, T. G., Lund, J.: The Short Rate Diffusion Revisited: A Tour Guided by the Efficient Method of Moments. Working Paper, 1996.

[5] Ayres, H. R., Barry, I. Y.: The Equilibrium Yield Curve for Government Securities. Financial Analysts Journal, 1979, 31-39.

[6] Ayres, H. R., Barry, I. Y.: A Theory of the U.S. Treasure Market Equilibrium. Management Science 26, 1980, 539-569.

[7] Ait-Sahalia, Y.: Transition Densities for Interest Rate and Other Nonlinear Diffusions. Journal of Finance 54, 1999, 1361-1395.

[8] Balduzzi, P., Das, S. R., Foresi, S.: The Central Tendency: A Second Factor in Bond Yields. Review of Economics and Statistics 80, 1998, 60-72.

[9] Barome-Adesi, G., Allegretto, A., Dinenis, E., Sorwar, G.: New numerical methods for the evaluation of interest rate contingent claims, Centre for Mathematical Trading and Finance Technical Report, City University Business School, 1997

[10] Brandt, M., Chua, C. T., Ramaswamy, K., Yaron, A.: A Two-Step Recursive Parameter Estimation Technique Using Cross-Sectional and Time-Series Data. Working paper, 2004.

[11] Brennan, M. J., Schwartz, E. S.: An Equilibrium Model of Bond Pricing and a Test of Market Efficiency. The Journal of Financial and Quantitative Analysis 17, 1982, 301-329.

[12] D. Brigo, F. Mercurio: Interest rate models - Theory and practice. With smile, inflation and credit. Springer, 2006.

[13] Buhler, W., Uhrig-Homburg, M., Walter, U., Weber, T.: An Empirical Comparison of Forward-Rate and Spot-Rate Models for Valuing Interest Rate Options. Journal of Finance 54, 1999, 269-305.

[14] Byers, S. L., Nowman, K. B.: Forecasting U.K. and U.S. Interest Rates Using Continuous Time Term Structure Models. International Review of Financial Analysis 7 (3), 1998, 191-206.

[15] Chan, K. C., Karolyi, G. A., Longstaff, F. A., Sanders, A. B.: An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. The Journal of Finance 47, 1992, 1209-1227.

[16] Chen, L.: Stochastic Mean and Stochastic Volatility - A Three-Factor Model of Term Structure of Interest Rates and its Applications to the Pricing of Interest Rate Derivatives. Blackwell Publishers, Oxford, 1996.

[17] Choi, Y., Wirjanto, T. S.: An analytic approximation formula for pricing zero coupon bonds. Finance Research Letters 4 (2), 2007, 116-126.

[18] Courtadon, G.: The pricing options on default-free bonds. Journal of Financial and Quantitative Analysis 17, 1982, 75-100.

[19] Cox, J. C., Ingersoll, J. E., Ross, S. A.: A Theory of the Term Structure of Interest Rates. Econometrica 53, 1985, 385-408.

[20] Christiansen, C.: Multivariate term structure models with levels and heteroskedasticity effects. Centre for analytical finance, University of Aarhus Working paper series, No. 129, 2002.

[21] Corzo, T., Schwartz, E. S.: Convergence within the European Union: evidence from Interest Rates, Economic Notes, 29(2), 2000, 243-268.

[22] Dell'Aquilla, R., Ronchetti, E., Trojani, F.: Robust GMM analysis for the short rate process. Journal of Empirical Finance 10, 2003, 373-397.

[23] Episcopos, A.: Further evidence on alternative continuous time models of the short term interest rate. Journal of International Financial Markets, Institutions and Money 10, 2000, 199-212.

[24] Fong, H. G., Vasicek, O. A.: Fixed-Income Volatility Management. Journal of Portfolio Management, 1991, 41-46.

[25] Fouque, J.-P. , Papanicolaou, G., Sircar, K. R.: Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press, 2000.

[26] Glasserman, P.: Monte Carlo Methods on Financial Engineering. Springer, 2004.

[27] Goodman, J., Moon, K. S., Szepessy, A., Tempone, R., Zouraris, G.: Stochastic and partial differential equations with adapted numerics. Royal Institute of Technology, Stockholm. www.math.kth.se/szepessy/sdepde.pdf

[28] Jensen, M. B.: Efficient Method of Moments Estimation of the Longstaff and Schwartz Interest rate Model. Centre for analytical finance, University of Aarhus Working paper series, No. 52, 2000.

[29] Kwok, Y. K.: Mathematical Models of Financial Derivatives. Springer, 1998.

[30] Longstaff, F. A., Schwartz, E. S.: Implementation of the Longstaff-Schwartz Interest Rate Model. Journal of Fixed Income 3, 1993, 7-14.

[31] Longstaff, F. A., Schwartz, E. S.: Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model. Journal of Finance 47, 1992, 1259-1282.

[32] Nowman, K. B.: Gaussian estimation of single-factor continuous time models of the term structure of interest rates. Journal of Finance 52, 1997, 1695-1706.

[33] Nowman, K. B.: The volatility of Japanese interest rates: evidence for Certificate of Deposit and Gensaki rates. International Review of Financial Analysis 11 (1), 2002, 29-38.

[34] Nowman, K. B., Sorwar, G.: Computation of Japanese bonds and derivative securities. Mathematics and Computers in Simulations 47, 1998, 583-588.

[35] Nowman, K. B., Sorwar, G.: Pricing UK and US securities within the CKLS model: Further results. International Review of Financial Analysis 8 (3), 1999, 235-245.

[36] Nowman, K. B., Sorwar, G.: Derivatives prices from interest rate models: results for Canada, Hong Kong and United States. International Review of Financial Analysis 14, 2005, 428-438.

[37] Oksendal, B.: Stochastic Differential Equations: An Introduction with Applications. Springer, 2003.

[38] Santamaria, T. C., Biscarri, J. G. : Nonparametric Estimation of Convergence of Interest Rates: Effects on Bond Pricing. Spanish Economic Review 7 (3), 2005, 167-190.

[39] Selby, M. J. P., Strickland, C.: Computing the Fong and Vasicek Pure Discount Bond Formula. Journal of Fixed Income, 1995, 78-84.

[40] Schaefer, S. M., Schwartz, E. S.: A Two-Factor Model of the Term Structure: An Approximate Analytical Solution. The Journal of Financial and Quantitative Analysis 19, 1984, 413-424.

[41] Sorwar, G., Barone-Adesi, G., Allegretto, W.: Valuation of derivatives based on single-factor interest rate models. Global Finance Journal 18, 2007, 251-269.

[42] Stehlíková, B., Ševčovič, D.: On a Volatility Averaging in a Two-Factor Interest Rate Model. Proceedings of Algoritmy 2005, Eds. A. Handlovičová, Z. Krivá, K. Mikula, D. Ševčovič, 2005, 325-333.

[43] Stehlíková, B.: Modeling Volatility Clusters with Application to Two-Factor Interest Rate Models. Journal of Electrical Engineering 56 (12/s), 2005, 90-93.

[44] Stehlíková, B.: Fast Mean Reverting Volatility in Fong-Vasicek Model of Interest Rates. Journal of Electrical Engineering 57 (12/s), 2006, 65-97.

[45] Stehlíková, B.: Averaged Bond Prices in Generalized Cox-Ingersoll-Ross Model of Interest Rates. Proceedings of 5th Actuarial and Financial Mathematics Day, Eds. M. Vanmaele, G. Deelstra, A. De Schepper, J. Dhaene, H. Reynaerts, W. Schoutens, P. Van Goethem, Koninklijke Vlaamse Acadamie van Belgie voor Wetenschappen en Kunsten, 2007, 77-87.

[46] Stehlíková, B.: Averaged Bond Prices for Fong-Vasicek and the Generalized Vasicek Interest Rates Models. Proceeding of MMEI, Eds. K. Cechlárová, M. Halická, V. Borbeľová, V. Lacko, 2007, 166-175.

[47] Stehlíková, B., Ševčovič, D.: Approximate formulae for pricing zero coupon bonds and their asymptotic analysis. To appear in: International Journal of Numerical Mathematics and Modeling.

[48] Ševčovič, D., Urbánová Csajková, A.: Calibration of one factor interest rate models, Journal of Electrical Engineering 55 (12/s), 2004, 46-50.

[49] Ševčovič, D., 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 Operation Research 13, 2005, 169-188.

[50] Treeponkaruna, S., Grey, S.: On the robustness of short rate models. Accounting and Finance 43, 2003, 87-121.

[51] Vasicek, O. A.: An Equilibrium Characterization of the Term Structure. Journal of Financial Economics 5, 1977, 177-188.

[52] Yu, J., Phillips, P. C. B.: Gaussian Estimation of Continuous Time Models of the Short Term Interest Rate. Cowles Foundation discussion paper no. 1309. Cowles Foundation for Research in Economics, Yale University, 2001.