Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics
Department of Applied Mathematics and Statistics

Two-factor Convergence Model of Cox-Ingersoll-Ross Type

Master's Thesis

Vladimír LACKO

9.1.9 Applied Mathematics
Economic and Financial Mathematics

Supervisor:
Beáta STEHLÍKOVÁ, PhD.

BRATISLAVA 2010

Two-factor Convergence Model of Cox-Ingersoll-Ross Type

Vladimír LACKO
E-mail: *lackovladimir@gmail.com*
Website: *http://www.st.fmph.uniba.sk/~lacko10*

Beáta STEHLÍKOVÁ
E-mail: *stehlikova@pc2.iam.fmph.uniba.sk*
Website: *http://www.iam.fmph.uniba.sk/institute/stehlikova*

Department of Applied Mathematics and Statistics
Faculty of Mathematics, Physics and Informatics
Comenius University in Bratislava
842 48 Bratislava
Slovakia

©2010 Vladimír Lacko and Beáta Stehlíkova
Master's thesis in Applied Mathematics
Compilation date: April 21, 2010
Typeset in LaTeX

Abstract

Corzo and Schwartz [2000, Convergence within the European Union: Evidence from Interest Rates, *Economic Notes* **29**, pp. 243–268] proposed a short-rate model for a country before adopting the Euro currency, which is based on the Vasicek model. In the first part of this work we provide a correct solution of the Corzo and Schwartz model and study an analogous model with the Cox-Ingersoll-Ross model applied. We show that the separation of the bond price can be done only in the case of uncorrelated increments of Wiener processes in stochastic differential equations for the European and domestic rates. Taking the bond price for an uncorrelated case as an approximation of a case with a correlation, we show that the difference between the logarithm of the bond price with and without a correlation is of the third order with respect to the time of maturity. In the second part of this work we propose a simple method for estimating parameters and compare both convergence models.

Keywords: two-factor • term structure • short-rate • convergence model • Vasicek • CIR • zero-coupon • bond • approximation • order of accuracy.

AMS Subject Classification: 91B28 • 91B70 • 60H10 • 35K10.

Acknowledgement

I am deeply indebted to my supervisor Dr. Beáta Stehlíková for devoting her enthusiasm and time to our co-operation, which resulted in this work. I also thank prof. Daniel Ševčovič and prof. Pavel Brunovský for their valuable comments, and Dr. Debra Gambrill for the grammar correction of this text. I am grateful to my parents for supporting me during my studies.

Declaration on Word of Honour

I declare on my honour that this work is based only on my knowledge, references and consultations with my supervisor.

....
Vladimír LACKO

UNIVERZITA KOMENSKÉHO V BRATISLAVE
Fakulta matematiky, fyziky a informatiky
Katedra aplikovanej matematiky a štatistiky

DVOJFAKTOROVÝ KONVERGENČNÝ MODEL TYPU COX-INGERSOLL-ROSS

Diplomová práca

Bc. Vladimír LACKO

9.1.9 Aplikovaná matematika
Ekonomická a finančná matematika

Školiteľ:
RNDr. Beáta STEHLÍKOVÁ, PhD.

BRATISLAVA 2010

Abstrakt

Corzová a Schwartz [2000, Convergence within the European Union: Evidence from Interest Rates, *Economic Notes* **29**, pp. 243–268] navrhli model okamžitej úrokovej miery pre krajiny, ktoré majú prijať menu Euro. Uvedený model je založený na Vašíčkovom modeli úrokovej miery. V prvej časti práce podávame opravené riešenie modelu Corzovej a Schwartza a skúmame analogický model aplikovaním Cox-Ingersoll-Rossovho modelu. Ukážeme, že separácia ceny dlhopisu (ako je tomu v článku Corzovej a Schwartza) je možná iba v prípade nekorelovaných prírastkov Wienerových procesov v stochastických diferenciálnych rovniciach pre domáci a európsky úrok. Ďalej ukážeme, že ak vezmeme cenu dlhopisu v prípade bez korelácie prírastkov Wienerových procesov ako aproximáciu riešenia s korelovanými prírastkami Wienerových procesov, potom rozdiel logaritmov ceny dlhopisov s koreláciou a bez nej je tretieho rádu vzhľadom na čas do maturity. V druhej časti práce navrhujeme jednoduchú metódu na odhad parametrov modelu a oba konvergenčné modely porovnávame.

Kľúčové slová: dvojfaktorový • konvergenčný model • časová štruktúra • úroková miera • Vašíček • CIR • bezkupónový • dlhopis • aproximácia • rád presnosti.

Contents

Introduction

Like music or art, mathematical equations can have a natural progression and logic that can evoke rare passions in a scientist. Although the lay public considers mathematical equations to be rather opaque, to a scientist an equation is very much like a movement in a larger symphony.

Michio Kaku

A *bond* is one of the most basic and widespread financial instruments. The Slovak Act No. 530/1990 Coll. on Bonds (*www.nbs.sk*) says

> A bond shall be a security with which is connected the right of the holder to require repayment of a sum owed in a nominal amount (*the par value*) and payment of yields on it (*coupons*) at a certain date (*the maturity*) and the duty of the person authorised to issue (the issuer) the bonds to fulfil these obligations.

There are different types of bonds, for instance, a fixed rate bond, an inflation-linked bond, a floating rates bond etc. The simplest type of bond is the so-called *zero-coupon bond*: a contract to repay borrowed money (*the principal*) with interest at the maturity date. Specifically, a zero-coupon bond with a unit par value is called a *discount bond*. The interest is usually determined by the *interest rate*.

This thesis deals with the question "How much should the bond cost?" or better "How much should the interest rate be?" It is clear that the price of a discount bond is given by

$$P(r_t, t, T) = \exp\{-R(r_t, t, T)(T - t)\},$$

where r_t is a *short-rate* at the time t, and $R(r_t, t, T)$ is a continuous interest rate for the period (t, T). The evolution of the short-rate is related to many factors, for instance, economic growth, crises, politics, etc. Once you buy a bond, you lose the possibility to make another investment. If the interest rate starts to increase, bonds become cheaper and you take a loss, and vice versa. It is a natural expectation that the price of a bond should be chosen in a way that one side cannot take advantage of the other. One approach, which we discuss in this thesis, is to model the evolution of the interest rate using stochastic processes. Once we have a model for the evolution of the interest rate and assume there are not any arbitrage opportunities we can compute "neutral" interest rates for any interval (t, T) which is represented by the term structure of interest rate $R(r_t, t, T)$.

The thesis is divided into four chapters. In Chapter 1 we deal with the basic stochastic calculus and processes, and we introduce some well-known term structure models. The second chapter, which contains the main theoretical results, is dedicated to a newly proposed convergence model. Chapter 3 is focused on a practical part of the thesis: a calibration of convergence models. In the last chapter, Chapter 4, we summarize the main results of this thesis and offer possibilities for further research.

Main goals of the thesis

The main goals of this thesis can be stated in the following three points:

- To find a correct solution to the convergence model by Corzo and Schwartz (2000), and state and prove new properties of this model.
- Formulate a new convergence model, find the corresponding bond price, and state and prove its properties.
- Propose an estimation method for convergence models, and compare the new model with the convergence model formulated by Corzo and Schwartz (2000) as well as with a few well known one-factor models.

List of symbols and abbreviations

$\boldsymbol{x}, \boldsymbol{y}$	vectors
$\boldsymbol{A}'$	transposition of $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	matrix inverse to $\boldsymbol{A}$
$\exp\{x\}$	e^x
$f(.)$	function
$\boldsymbol{f}(.)$	vector function
$\boldsymbol{f}^{\mathrm{inv}}(.)$	inverse transformation to $\boldsymbol{f}$
$\boldsymbol{\nabla}_{\boldsymbol{x}} f$	gradient of a scalar function f with respect to $\boldsymbol{x}$
$\boldsymbol{\nabla}^2_{\boldsymbol{x}} f$	Hessian matrix of a scalar function f with respect to $\boldsymbol{x}$
$\mathbf{J}_{\boldsymbol{x}} \boldsymbol{f}$	Jacobian matrix of a vector function $\boldsymbol{f}$ with respect to $\boldsymbol{x}$
$\mathrm{d}\boldsymbol{f}(t), \mathrm{d}\boldsymbol{f}_t, \mathrm{d}\boldsymbol{f}$	differential of $\boldsymbol{f}$ with respect to t
$\dot{\boldsymbol{f}}(t), \dot{\boldsymbol{f}}, \mathrm{d}_t \boldsymbol{f}$	derivative of $\boldsymbol{f}$ with respect to t
$\partial_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{y})$	partial derivative of a scalar function f with respect to $\boldsymbol{x}$

$\partial^2_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{y})$	second partial derivative of f with respect to $\boldsymbol{x}$
$\partial^2_{\boldsymbol{x}\boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y})$	second partial derivative of f with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$
$\boldsymbol{X}, \boldsymbol{Y}$	random vectors
$\boldsymbol{X}(t, \omega), \boldsymbol{X}(t), \boldsymbol{X}_t$	t-parametrized random vector
$\mathrm{Cov}[X, Y]$	covariance of random variables X and Y
$\mathrm{Cor}[X, Y]$	correlation of random variables X and Y
$X \mid Y$	random variable X conditioned on Y
$\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$	normal distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$
ODE	ordinary differential equation
PDE	partial differential equation
SDE	stochastic differential equation
EMU	European Monetary Union

1

A brief theory of bond-pricing models

1.1 Term structures of interest rates

Definition 1. *A bond is a debt security by which the authorized issuer owes the holders a debt and is obliged to pay interest (the coupon) and/or to repay the principal at a later day, termed the maturity. If there is no coupon payment, the bond is called a zero-coupon. The face value of the bond is called its par value. A zero-coupon bond with a par value of* 1 *is called a discount bond.*

Let r_t be a *short-rate* at the time t and $R(r_t, t, T)$ be a *continuous interest rate* for the period (t, T) of the length $\tau = T - t$. Then the price of the discount bond is given by

$$P(r_t, t, T) = \exp\{-R(r_t, t, T)(T - t)\}.$$

After some transformation we obtain that

$$R(r_t, t, T) = -\frac{\ln[P(r_t, t, T)]}{(T - t)} = -\frac{\ln[P(r_t, \tau)]}{\tau}.$$

The function $R(r, t, T)$ is the so-called *term structure of interest rates* or a *yield curve*. Figure 1.1 depicts an example of the term structures of EURIBOR, LIBOR, PRIBOR and BRIBOR.

Given the prices of discount bonds we are able to figure out the value of the short-rate by r_t. It is clear that $r_t = R(r_t, t, t)$, and

$$R(r, t, t) = \lim_{T \to t} \left(-\frac{\ln[P(r, t, T)] - \ln[P(r, t, t)]}{(T - t)} \right) = -\partial_T \ln[P(r, t, T)] \Big|_{T=t}.$$

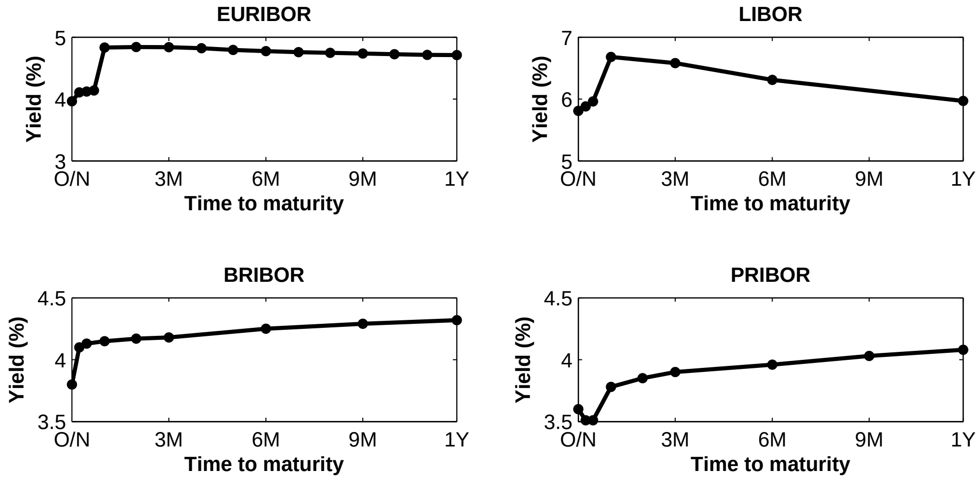

Figure 1.1: *European (EURIBOR), British (LIBOR), Slovak (BRIBOR) and Czech (PRIBOR) term structures of interbank interest rates on* 3^{rd} *December 2007. Source: www.euribor.org, http://www.bankofengland.co.uk, http://www.nbs.sk and http//www.cnb.cz.*

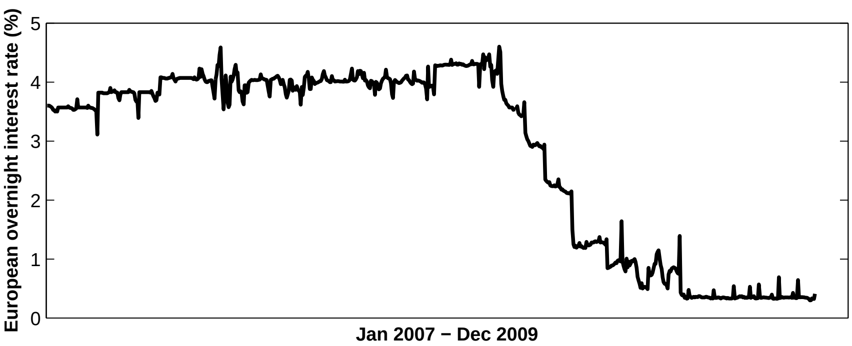

Figure 1.2: *European overnight interest rate.*

1.2 Stochastic processes and Itō's lemma

It is an undisputed fact that the evolution of an interest rate has a stochastic character. Figure 1.2 demonstrates the EMU's overnight interest rate (EONIA) from January 2007 to December 2009. We can see that the process behaves like a sequence of random variables that form a fractal curve. Such a curve is not intuitively differentiable; consequently, the ordinary calculus is not applicable. An important tool in modelling interest rates and pricing bonds is the probability theory. In particular, stochastic processes, stochastic differential equations and Itō's lemma seem to be a suitable way to describe and work with interest rates. The aim of this section is to discuss the basic notions.

Since we work with randomness and uncertainty, a probability space is "the place to work." A probability space is a triple $(\Omega, \mathcal{F}, \Pr)$, where Ω is a given set of all elementary events, $\mathcal{F}$ is a σ-algebra (a nonempty collection of subsets of Ω

(including Ω itself) that is closed under complementation and countable unions of its members), and $\Pr$ is a (measurable) function such that $\Pr : \mathcal{F} \to [0, 1]$. We do not discuss all the basics of the probability theory in this section. The reader is referred to Øksendal (2000) for details.

Stochastic processes

In the following two definitions we frame two important concepts: a stochastic process and its special case, a Wiener process.

Definition 2. *A stochastic process is a parametrized collection of random variables (vectors)* $\{\boldsymbol{X}(t, \omega)\}_{t \in \mathcal{T}}$*, defined on a probability space* $(\Omega, \mathcal{F}, \Pr)$ *with values in* $\mathbb{R}^n$*.*

The set $\mathcal{T}$ is usually an interval $[0, T]$ or a halfline $[0, \infty)$. We offer two views of a stochastic process: Firstly, a function $\omega \to \boldsymbol{X}(t, \omega)$, t fixed and $\omega \in \Omega$, is a random variable. That is, at each time t we obtain a realization of the random variable $\boldsymbol{X}$. Secondly, for a fixed $\omega \in \Omega$ we obtain the function $t \to \boldsymbol{X}(t, \omega)$, which is called a path of $\boldsymbol{X}(t, \omega)$. We can represent these two points of view as follows: one particular pollen seed in a water takes one particular path, and different seeds take different paths.

The most important special case is a *Wiener process*, which we describe in the following definition

Definition 3. *A (one-dimensional) Wiener process* $\{W(t),\, t \geq 0\}$ *is a continuous-time stochastic process characterized by three facts:* ***i)*** *all increments* $W(t + \Delta) - W(t) \sim \mathcal{N}(0, \Delta)$*;* ***ii)*** *for each time partitioning* $t_0 = 0 < t_1 < t_2 < \ldots < t_n$ *increments* $W(t_1) - W(t_0),\, W(t_2) - W(t_1), \ldots,\, W(t_n) - W(t_{n-1})$ *are mutually independent;* ***iii)*** $\Pr[W(0) = 0] = 1$.

Figure 1.3 illustrates a few paths of a one-dimensional Wiener process. Each path corresponds to one "seed."

We denote the increment $W(t + \mathrm{d}t) - W(t)$ of a Wiener process by $\mathrm{d}W(t)$. We can also define an n-dimensional Wiener process

$$\{\boldsymbol{W}(t),\, t \geq 0\} = \{(W_1(t), \ldots, W_n(t))',\, t \geq 0\},$$

where $W_i(t)$, $i = 1, \ldots, n$, are independent one-dimensional Wiener processes.

Stochastic differentials and Itō's lemma

Let us introduce a stochastic differential equation using an example in economics. Suppose a constant continuous interest rate r and an initial zero-coupon bond price P_0. Then the price of a bond at the time t is $P(r, t) = P_0 \exp\{rt\}$, or, in the words of ordinary differential equations (ODEs), $\dot{P} = rP$ with an initial condition $P(0) = P_0$. Now, assume that the interest rate is not constantly equal to r but is a deterministic time-dependent function, i.e., $r = r(t)$. Then the bond price satisfies ODE $\dot{P} = r(t)P$, and its solution is $P(r(t), t) = P_0 \exp\{\int_0^t r(s)\mathrm{d}s\}$. Finally, there might

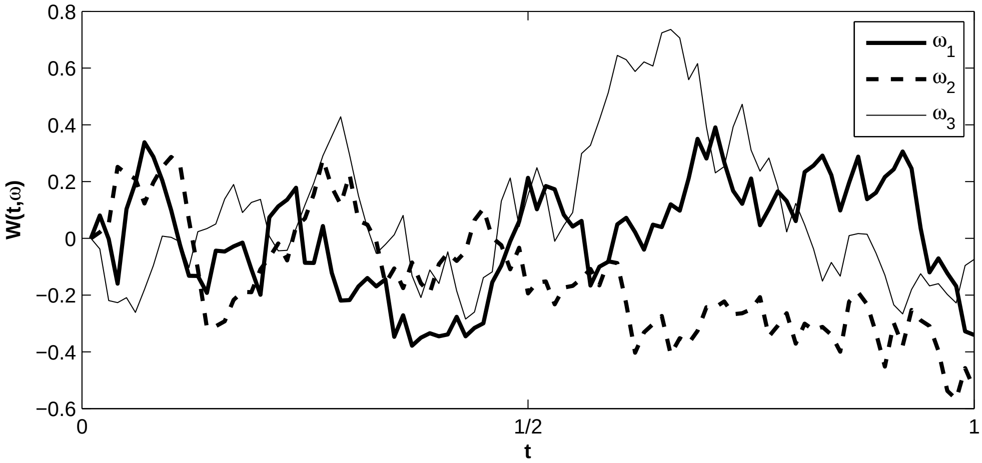

Figure 1.3: *A few simulated paths of a one-dimensional Wiener process.*

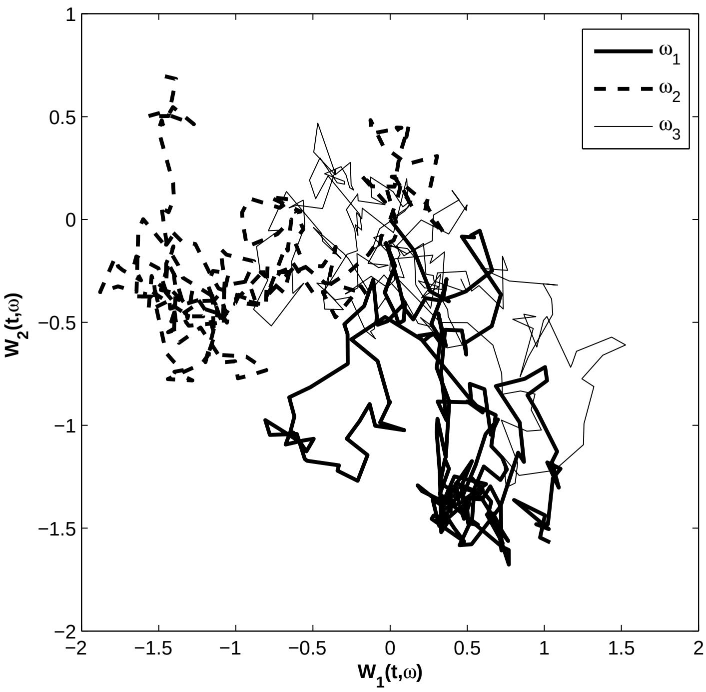

Figure 1.4: *A few simulated paths of a two-dimensional Wiener process with uncorrelated components of the increments.*

also be some random fluctuations in the evolution of r, i.e., $r = r(t) +$ "noise". In this case we obtain that $\dot{P} = (r(t) + \text{"noise"})P$, that is, a differential equation with a stochastic element, the so-called *stochastic differential equation* (SDE). However, bond-pricing is not that easy, since there are more complicated models for describing the evolution of interest rates.

Usual ways to model the evolution of interest rates are stochastic differential equations of the form

$$\mathrm{d}\boldsymbol{X} = \boldsymbol{\mu}(\boldsymbol{X}, t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}, t)\mathrm{d}\boldsymbol{W}, \tag{1.1}$$

and its solution is a stochastic process $\boldsymbol{X}$. To simplify the notation we use $\boldsymbol{X}$ instead of $\boldsymbol{X}(t)$ and $\mathrm{d}\boldsymbol{W}$ instead of $\mathrm{d}\boldsymbol{W}(t)$. It is desirable to know the distribution of $\boldsymbol{X}$ at each time t. Equation (1.1) is also called a *diffusion* or an *Itō process*. The component-wise form of the equation (1.1) is

$$\begin{aligned}
\mathrm{d}X_1 &= \mu_1(\boldsymbol{X}, t)\mathrm{d}t + \sigma_{11}(\boldsymbol{X}, t)\mathrm{d}W_1 + \cdots + \sigma_{1m}(\boldsymbol{X}, t)\mathrm{d}W_m, \\
&\vdots \\
\mathrm{d}X_n &= \mu_n(\boldsymbol{X}, t)\mathrm{d}t + \sigma_{n1}(\boldsymbol{X}, t)\mathrm{d}W_1 + \cdots + \sigma_{nm}(\boldsymbol{X}, t)\mathrm{d}W_m.
\end{aligned}$$

Note that $\boldsymbol{X}$ is a *Markovian process*, since the increment only depends on the present value of $\boldsymbol{X}$.

However, we are not only interested in the value of an interest rate described by a SDE, but also in the bond price, which is a function of the interest rate. How to deal with this kind of problem is described in one of the most famous lemmas:

Lemma 4 (Theorem 6 in Itō (1951) – Itō's lemma). *Let* $f(\boldsymbol{x}, t) : \mathbb{R}^n \times \langle 0, \infty) \to \mathbb{R}$, $f \in \mathcal{C}^2$, $\boldsymbol{\mu}(\boldsymbol{x}, t) : \mathbb{R}^n \times \langle 0, \infty) \to \mathbb{R}^n$, $\boldsymbol{\sigma}(\boldsymbol{x}, t) : \mathbb{R}^n \times \langle 0, \infty) \to \mathbb{R}^{n \times m}$, $\boldsymbol{W}$ *be an* m*-dimensional Wiener process, and let* $\boldsymbol{X}$ *be a stochastic process satisfying* (1.1). *Then the process* $f(\boldsymbol{X}, t)$ *satisfies*

$$\mathrm{d}f = \partial_t f \mathrm{d}t + (\boldsymbol{\nabla}_{\boldsymbol{x}} f)' \mathrm{d}\boldsymbol{X} + \frac{1}{2}(\mathrm{d}\boldsymbol{X})'(\boldsymbol{\nabla}^2_{\boldsymbol{x}} f)\mathrm{d}\boldsymbol{X},$$

where $\mathrm{d}W_i \mathrm{d}W_j = \delta_{ij}\mathrm{d}t$, *and* $\mathrm{d}W_i \mathrm{d}t = \mathrm{d}t\mathrm{d}W_i = 0$.

We note that if $\mathrm{Cor}[\mathrm{d}W_1, \mathrm{d}W_2] = \rho$, then $\mathrm{d}W_1 \mathrm{d}W_2 = \rho \mathrm{d}t$ in Itō's lemma.

1.3 One-factor models

In the previous section we introduced the basic mathematical tools needed for modelling interest rates. *One-factor interest rate models* describe a change in the value of an interest rate $\mathrm{d}r$ depending on only one factor: r itself. That is, we can characterize the evolution of an interest rate by one SDE of the form

$$\mathrm{d}r = \mu(r, t)\mathrm{d}t + \sigma(r, t)\mathrm{d}W. \tag{1.2}$$

The function $\mu(r, t)$ is the so-called *drift*, and the function $\sigma(r, t)$ is the so-called *volatility* or diffusion. The choice of functions μ and σ gives different one-factor short-rate models. In the following we describe a few well-known one factor models.

Examples of one-factor models

Vasicek (1977) proposed a simple mean-reversion model with a constant volatility:

$$\mathrm{d}r = \kappa(\theta - r)\mathrm{d}t + \sigma \mathrm{d}W, \tag{1.3}$$

where $\kappa,\ \sigma > 0$ and $\theta \geq 0$. This model is also referred as an *Ornstein-Uhlenbeck mean-reversion*. A disadvantage of this model is that r may reach negative values, although interest rates should only reach positive values.

Cox, Ingersoll, Jr. and Ross (1985) (CIR) replaced the constant volatility in the Vasicek model by its $\sqrt{r}$ multiple. More precisely, they stated their model as

$$\mathrm{d}r = \kappa(\theta - r)\mathrm{d}t + \sigma\sqrt{r}\mathrm{d}W, \tag{1.4}$$

which is also called a *Bessel square root mean-reversion process*. In the CIR model the volatility decreases by decreasing r; therefore, the process cannot reach a negative value. Due to Itō's lemma, the process $x = \ln(r)$ satisfies (Kwok (2008))

$$\mathrm{d}x = (2\mathrm{e}^{-x}(2\kappa\theta - \sigma^2) - \kappa)\mathrm{d}t + \mathrm{e}^{-x/2}\mathrm{d}W.$$

If $r = 0$ (the stochastic element is zero, and only the deterministic part remains), then $x \to -\infty$. The condition $2\kappa\theta \geq \sigma^2$ ensures that the drift $2\mathrm{e}^{-x}(2\kappa\theta - \sigma^2) - \kappa \to \infty$ for $x \to -\infty$ (that is, the more r gets closer to 0, the more $\mathrm{d}x$ increases). This eliminates the possibility of $x \to -\infty$ (that is $r \to 0$); consequently, the probability of a non-positive interest rate is zero. If $2\kappa\theta < \sigma^2$, then for $x \to -\infty$ the drift $2\mathrm{e}^{-x}(2\kappa\theta - \sigma^2) - \kappa \to -\infty$, hence $r \to 0$ faster.

A generalization of the previous two models was proposed by Chan, Karolyi, Longstaff and Sanders (1992), the *CKLS model*

$$\mathrm{d}r = \kappa(\theta - r)\mathrm{d}t + \sigma r^{\gamma}\mathrm{d}W,$$

where $\gamma \geq 0$. It was shown that γ is not necessarily equal to 0 (the Vasicek model) or $1/2$ (the CIR model). Chan et al. (1992) estimated the general model and its versions using U.S. Treasury bill yields. They also reported that γ is usually greater than 1 in an unconstrained estimation.

Bond-pricing partial differential equation for a one-factor model

In the following we derive the *bond-pricing partial differential equation* (PDE) (see, e.g., Kwok (1998)). Let r follow SDE (1.2), and $P(r, t, T)$ be the price of a discount bond. Then, Itō's lemma implies that P satisfies

$$\mathrm{d}P = \left(\partial_t P + \mu\partial_r P + \frac{1}{2}\sigma^2\partial_r^2 P\right)\mathrm{d}t + \sigma\partial_r P\mathrm{d}W = \mu_P\mathrm{d}t + \sigma_P\mathrm{d}W,$$

where we have denoted $\mu_P = \partial_t P + \mu\partial_r P + \frac{1}{2}\sigma^2\partial_r^2 P$ and $\sigma_P = \sigma\partial_r P$.

Let us assume the following portfolio: 1 bond with a maturity T_1 and Δ bonds with a maturity T_2. Then the value Π of the portfolio is $\Pi = P(r, t, T_1) + \Delta P(r, t, T_2)$, and the change in the value of the portfolio is $\mathrm{d}\Pi = \mathrm{d}P(r,t,T_1) + \Delta P(r,t,T_2)$. Setting SDEs for both bonds into the equation for $\mathrm{d}\Pi$ yields

$$\mathrm{d}\Pi = (\mu_P(r,t,T_1) + \Delta\mu_P(r,t,T_2))\mathrm{d}t + (\sigma_P(r,t,T_1) + \Delta\sigma_P(r,t,T_2))\mathrm{d}W.$$

By setting $\Delta = -\sigma_P(r,t,T_1)/\sigma_P(r,t,T_2)$ into the previous equation, we eliminate the stochastic term, that is

$$\mathrm{d}\Pi = \left(\mu_P(r,t,T_1) - \frac{\sigma_P(r,t,T_1)}{\sigma_P(r,t,T_2)}\mu_P(r,t,T_2)\right)\mathrm{d}t.$$

Since we rule out any arbitrage opportunities, the right-hand side must be equal to $r\Pi\mathrm{d}t$. We obtain that

$$\mu_P(r,t,T_1) - \frac{\sigma_P(r,t,T_1)}{\sigma_P(r,t,T_2)}\mu_P(r,t,T_2) = r\left(P(r,t,T_1) - \frac{\sigma_P(r,t,T_1)}{\sigma_P(r,t,T_2)}P(r,t,T_2)\right),$$

which, after some transformation, implies that

$$\frac{\mu_P(r,t,T_1) - rP(r,t,T_1)}{\sigma_P(r,t,T_1)} = \frac{\mu_P(r,t,T_2) - rP(r,t,T_2)}{\sigma_P(r,t,T_2)}$$

The previous equality holds for any T_1 and T_2; therefore, the ratio does not depend on the time to maturity. We define

$$\lambda(r,t) = \frac{\mu_P(r,t,T) - rP(r,t,T)}{\sigma_P(r,t,T)},$$

where λ is the so-called *market price of risk*. Setting μ_P and σ_P into the previous equation yields that the price of the discount bond must satisfy PDE

$$\partial_t P + (\mu - \lambda\sigma)\partial_r P + \frac{1}{2}\sigma^2\partial_r^2 P - rP = 0, \tag{1.5}$$

with the terminal condition $P(r,T,T) = 1$. If functions μ, σ and λ only depend on the time to maturity $T-t$ and the interest rate r, the transformation $\tau = T-t$ only changes the first term to $-\partial_\tau P$ and, instead of the terminal condition, we have the initial condition $P(r,0) = 1$. After we have derived the bond-pricing PDE we can figure out prices of a discount bond in some of the previously mentioned models.

Example: Pricing a bond in the case of the Vasicek model

If we assume a constant market price of risk λ in the Vasicek model (see SDE (1.3)), we obtain that the price of a discount bond solves

$$-\partial_\tau P + [\kappa(\theta - r) - \lambda\sigma]\partial_r P + \frac{1}{2}\sigma^2\partial_r^2 P - rP = 0,\ P(r,0) = 1. \tag{1.6}$$

Let us consider a solution of the form

$$P(r,\tau) = \exp\{A(\tau) - rD(\tau)\}. \tag{1.7}$$

Evidently, the initial condition $P(r, 0) = 1$ for all r implies that $A(0) = D(0) = 0$. A solution of the form (1.7) gives us $\partial_\tau P = (\dot{A} - r\dot{D})P$, $\partial_r P = -DP$ and $\partial_r^2 P = D^2 P$. Therefore, setting this solution to (1.6) yields

$$-(\dot{A} - r\dot{D}) - [\kappa(\theta - r) - \lambda\sigma]B + \frac{1}{2}\sigma^2 B^2 - r = 0.$$

After some transformation we obtain that

$$r(\dot{D} + \kappa D - 1) + [-\dot{A} - (\kappa\theta - \lambda\sigma)D + \frac{1}{2}\sigma^2 D^2] = 0.$$

Since the previous equation holds for any r, the following system of ODEs have to be satisfied:

$$\begin{aligned} \dot{D} &= 1 - \kappa D, \\ \dot{A} &= -(\kappa\theta - \lambda\sigma)D + \frac{1}{2}\sigma^2 D^2, \\ & A(0) = 0,\ D(0) = 0. \end{aligned}$$

A solution to D is

$$D(\tau) = \frac{1 - \exp\{-\kappa\tau\}}{\kappa}.$$

A solution to A can be found by integration, that is

$$\begin{aligned} A(\tau) &= \int_0^\tau -(\kappa\theta - \lambda\sigma)D(s) + \frac{1}{2}\sigma^2 [D(s)]^2 \mathrm{d}s \\ &= \left[\frac{1 - \exp\{-\kappa\tau\}}{\kappa} - \tau\right] R_\infty - \frac{\sigma^2}{4\kappa^3}(1 - \exp\{-\kappa\tau\})^2, \end{aligned}$$

where $R_\infty = \theta - \lambda\sigma/\kappa - \sigma^2/(2\kappa^2)$, is the limit of the term structure for $\tau \to \infty$. We can see that the value of a short-rate does not influence the price of a discount bond with a long period of maturity.

Example: Pricing a bond in the case of the CIR model

Under the assumption that the market price of risk is $\lambda = \nu\sqrt{r}$, where ν is a constant, we obtain the bond-pricing PDE for the CIR model (see SDE (1.4))

$$-\partial_\tau P + [\kappa(\theta - r) - \nu\sigma r]\partial_r P + \frac{1}{2}\sigma^2 r \partial_r^2 P - rP = 0,\ P(r, 0) = 1. \quad (1.8)$$

Again, we consider a solution of the form (1.7). Setting such a solution to the previous PDE yields that A and D satisfy the following system of ODEs:

$$\begin{aligned} \dot{D} &= 1 - (\kappa + \nu\sigma)D - \frac{1}{2}\sigma^2 D^2, \\ \dot{A} &= -\kappa\theta D, \\ & A(0) = D(0) = 0. \end{aligned}$$

A solution to D (see the original paper by Cox et al. (1985)) is

$$D(\tau) = \frac{2(\exp\{\phi\tau\} - 1)}{(\phi + \psi)(\exp\{\phi\tau\} - 1) + 2\phi},$$

where $\psi = \kappa + \nu\sigma$ and $\phi = \sqrt{\psi^2 + 2\sigma^2} = \sqrt{(\kappa + \nu\sigma)^2 + 2\sigma^2}$. A solution to A can be found by integration, that is,

$$A(\tau) = \int_0^\tau -\kappa\theta D(s)\mathrm{d}s,$$

which implies

$$A(\tau) = \frac{2\kappa\theta}{\sigma^2} \ln\left[\frac{2\phi\exp\{(\phi + \psi)\tau/2\}}{(\phi + \psi)(\exp\{\phi\tau\} - 1) + 2\phi}\right].$$

We note that the CKLS model has no analytical solution for $\gamma \neq 0$ and $\gamma \neq 1/2$. Nevertheless, an analytical approximation was done by Choi and Wirjanto (2007).

1.4 Two-factor models

Two-factor models assume that the interest rate is a function of two factors (for instance, the sum of the factors or the value of one factor), where the dynamics of the factors is described by a system of SDEs. The usual approach is that the first factor is the interest rate, and the other governs a parameter in the first equation. We refer the reader to Kwok (1998) and Ševčovič et al. (2009) for more details on general two-factor models.

We consider the following two-factor model:

$$\begin{aligned} \mathrm{d}r &= \mu_\mathrm{r}(r,x,t)\mathrm{d}r + \sigma_\mathrm{r}(r,x,t)\mathrm{d}W_1, \\ \mathrm{d}x &= \mu_\mathrm{x}(r,x,t)\mathrm{d}x + \sigma_\mathrm{x}(r,x,t)\mathrm{d}W_2, \\ \mathrm{Cov}[\mathrm{d}W_1, \mathrm{d}W_2] &= \rho\mathrm{d}t. \end{aligned} \tag{1.9}$$

where x is the factor that influences the interest rate r, and ρ is a constant correlation between the increments $\mathrm{d}W_1$ and $\mathrm{d}W_2$ of the Wiener processes. We note that it is possible to rewrite the model (1.9) using the independent increments $\mathrm{d}\widetilde{W}_1$ and $\mathrm{d}\widetilde{W}_2$ as follows:

$$\begin{aligned} \mathrm{d}r &= \mu_\mathrm{r}(r,x,t)\mathrm{d}r + \sigma_\mathrm{r}(r,x,t)\mathrm{d}\widetilde{W}_1, \\ \mathrm{d}x &= \mu_\mathrm{x}(r,x,t)\mathrm{d}x + \sigma_\mathrm{x}(r,x,t)[\rho\mathrm{d}\widetilde{W}_1 + (1-\rho^2)^{1/2}\mathrm{d}\widetilde{W}_2]. \end{aligned} \tag{1.10}$$

Examples of two-factor models

A well-known class of two-factor models is the so-called *convergence model*. It describes the behaviour of the interest rate of a country entering a monetary union. The evidence is that the domestic rate converges to the union's rate and that both rates are strongly correlated. For instance, Figure 1.5, which depicts the Slovak (BRIBOR) and EMU's overnight (EONIA) interest rate before Slovakia adopted the

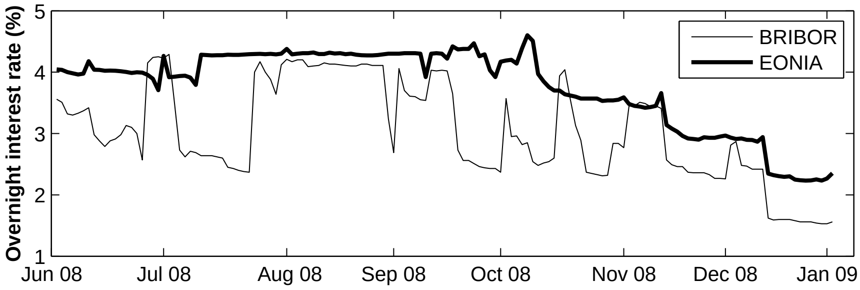

Figure 1.5: *An example of a convergence within the Slovak(BRIBOR) and the EMU's (EONIA) over-night interest rate before Slovakia adopted the Euro currency.*

Euro currency on 1^{st} January 2009, confirms that the Slovak interest rate converged to the EMU's rate, and both were correlated.

In a pioneering paper by Corzo and Schwartz (2000), the first convergence model was formulated; in particular, the convergence of the Spanish interest rate was discussed. The authors applied the Vasicek model, that is

$$\begin{aligned} \mathrm{d}r_{\mathrm{d}} &= [a + b(r_{\mathrm{u}} - r_{\mathrm{d}})]\mathrm{d}t + \sigma_{\mathrm{d}}\mathrm{d}W_{\mathrm{d}}, \\ \mathrm{d}r_{\mathrm{u}} &= c(d - r_{\mathrm{u}})\mathrm{d}t + \sigma_{\mathrm{u}}\mathrm{d}W_{\mathrm{u}}, \\ \mathrm{Cov}[\mathrm{d}W_{\mathrm{d}}, \mathrm{d}W_{\mathrm{u}}] &= \rho \mathrm{d}t, \end{aligned} \tag{1.11}$$

where r_{d} corresponds to the domestic interest rate, and r_{u} corresponds to the interest rate of the EMU. In this model we suppose that the constants b, c, σ_{d} and σ_{u} are positive, a and d are non-negative, and $0 < \rho < 1$. The constant a is interpreted as a minor divergence (the domestic rate does not exactly replicate the EMU's rate). In this work we call this model a *convergence model of Vasicek type*.

Another interesting class of two-factor models are models with stochastic parameters. Anderson and Lund (1996) proposed a *two-factor stochastic volatility model*, where the interest rate is modelled by the CKLS model, and the volatility is modelled by the so-called logarithmic Vasicek model. That is, the model takes the form

$$\begin{aligned} \mathrm{d}r &= \kappa_1(\theta - r)\mathrm{d}t + \sigma r^{\gamma}\mathrm{d}W_1, \\ \mathrm{d}\ln(\sigma^2) &= \kappa_2[\theta_2 - \ln(\sigma^2)]\mathrm{d}t + \xi \mathrm{d}W_2, \end{aligned}$$

where $\mathrm{d}W_1$ and $\mathrm{d}W_2$ are independent. Another model for stochastic volatility was proposed by Fong and Vasicek (1991). They modified the Vasicek model:

$$\begin{aligned} \mathrm{d}r &= \kappa_1(\theta - r)\mathrm{d}t + \sqrt{v}\mathrm{d}W_1, \\ \mathrm{d}v &= \kappa_2[\theta_2 - v]\mathrm{d}t + \sigma\sqrt{v}\mathrm{d}W_2, \\ \mathrm{Cov}[\mathrm{d}W_1, \mathrm{d}W_2] &= \rho \mathrm{d}t. \end{aligned}$$

For more details on the previous stochastic volatility models, we refer the reader to Section 7.1.7 in Kwok (1998). Another approach to modelling parameters was introduced in a paper by Balduzzi et al. (1998). They suggest modelling the limit of the interest rate. More precisely, they assumed a model of the following form:

$$\begin{aligned} \mathrm{d}r &= \kappa(\theta - r)\mathrm{d}t + \sigma_{\mathrm{r}}(r)\mathrm{d}W_1, \\ \mathrm{d}\theta &= \mu(\theta)\mathrm{d}t + \sigma_{\theta}(\theta)\mathrm{d}W_2. \end{aligned}$$

Bond-pricing partial differential equation to a two-factor model

After we have introduced some two-factor models, we derive the bond-pricing PDE. The idea is analogous to the case of one-factor models. Let r and x follow the system of SDEs (1.9). Then, Itō's lemma yields

$$\mathrm{d}P = \mu_P \mathrm{d}t + \sigma_{P\mathrm{r}} \mathrm{d}W_1 + \sigma_{P\mathrm{x}} \mathrm{d}W_2, \tag{1.12}$$

where

$$\begin{aligned} \mu_P &= \partial_t P + \mu_\mathrm{r} \partial_r P + \mu_\mathrm{x} \partial_x P + \frac{1}{2}\sigma_\mathrm{r} \partial_r^2 P + \frac{1}{2}\sigma_\mathrm{x} \partial_x^2 P + \rho \sigma_\mathrm{x} \sigma_\mathrm{r} \partial_{rx}^2 P, \\ \sigma_{P\mathrm{r}} &= \sigma_\mathrm{r} \partial_r P, \\ \sigma_{P\mathrm{x}} &= \sigma_\mathrm{x} \partial_x P. \end{aligned}$$

Again we use the bond hedging and no-arbitrage principle (see Kwok (1998)) for three bonds with maturities T_1, T_2 and T_3. We denote $P(r, x, t, T_i) = P(T_i)$ to shorten some expressions. The quantity of the corresponding bond is denoted by V_1, V_2 and V_3. It is obvious that the value of the portfolio is given by $\Pi = P(T_1)V_1 + P(T_2)V_2 + P(T_3)V_3$. Using (1.12) we obtain that the change in the value of the portfolio is given by

$$\begin{aligned} \mathrm{d}\Pi &= V_1 \mathrm{d}P(T_1) + V_2 \mathrm{d}P(T_2) + V_3 \mathrm{d}P(T_3) \\ &= [V_1 \mu_P(T_1) + V_2 \mu_P(T_2) + V_3 \mu_P(T_3)]\mathrm{d}t \\ &\quad [V_1 \sigma_{P\mathrm{r}}(T_1) + V_2 \sigma_{P\mathrm{r}}(T_2) + V_3 \sigma_{P\mathrm{r}}(T_3)]\mathrm{d}W_1 \\ &\quad [V_1 \sigma_{P\mathrm{x}}(T_1) + V_2 \sigma_{P\mathrm{x}}(T_2) + V_3 \sigma_{P\mathrm{x}}(T_3)]\mathrm{d}W_2. \end{aligned}$$

To eliminate all stochastic terms the equalities

$$V_1 \sigma_{P\mathrm{r}}(T_1) + V_2 \sigma_{P\mathrm{r}}(T_2) + V_3 \sigma_{P\mathrm{r}}(T_3) = 0, \tag{1.13}$$

$$V_1 \sigma_{P\mathrm{x}}(T_1) + V_2 \sigma_{P\mathrm{x}}(T_2) + V_3 \sigma_{P\mathrm{x}}(T_3) = 0, \tag{1.14}$$

have to be satisfied. The rule out of arbitrage gives us the condition

$$V_1 \mu_P(T_1) + V_2 \mu_P(T_2) + V_3 \mu_P(T_3) = r\Pi = r[P(T_1)V_1 + P(T_2)V_2 + P(T_3)V_3],$$

which implies

$$V_1[\mu_P(T_1) - rP(T_1)] + V_2[\mu_P(T_2) - rP(T_2)] + V_3[\mu_P(T_3) - rP(T_3)] = 0. \tag{1.15}$$

The system (1.13) – (1.15) has a non-trivial solution $(V_1, V_2, V_3)^T$ if and only if at least one equation of the system (1.13) – (1.15) is a linear combination of the two others. Equations (1.13) and (1.14) are independent (otherwise, the problem is reduced to a one-factor model); therefore, (1.15) is a linear combination of (1.13) and (1.14), and we obtain that

$$\mu_P(r, x, t, T_i) - rP(r, x, t, T_i) = \lambda_\mathrm{r}(r, x, t)\sigma_{P\mathrm{r}}(r, x, t, T_i) + \lambda_\mathrm{x}(r, x, t)\sigma_{P\mathrm{x}}(r, x, t, T_i), \quad i = 1, 2, 3.$$

Since the previous equation holds for any T_i, the market price of risk λ_r for the interest rate r and the market price of risk λ_{x} for the factor x do not depend on T. Setting μ_P, $\sigma_{P\mathrm{r}}$ and $\sigma_{P\mathrm{x}}$ into the previous equation yields that the bond pricing PDE for model (1.9) is

$$\partial_t P + (\mu_{\mathrm{r}} - \lambda_{\mathrm{r}}\sigma_{\mathrm{r}})\partial_r P + (\mu_{\mathrm{x}} - \lambda_{\mathrm{x}}\sigma_{\mathrm{x}})\partial_x P + \frac{1}{2}\sigma_{\mathrm{r}}^2\partial_r^2 P + \frac{1}{2}\sigma_{\mathrm{x}}^2\partial_x^2 P + \rho\sigma_{\mathrm{x}}\sigma_{\mathrm{r}}\partial_{rx}^2 P - rP = 0. \tag{1.16}$$

Example: Pricing a bond in the case of the convergence model of Vasicek type and its analysis

In this example we focus our attention on the model (1.11) by Corzo and Schwartz. The corresponding bond pricing PDE (using the transformation $\tau = T - t$) is

$$\begin{aligned} -\partial_\tau P + [a + b(r_{\mathrm{u}} - r_{\mathrm{d}}) - \lambda_{\mathrm{d}}\sigma_{\mathrm{d}}]\partial_{r_{\mathrm{d}}} P + [c(d - r_{\mathrm{u}}) - \lambda_{\mathrm{u}}\sigma_{\mathrm{u}}]\partial_{r_{\mathrm{u}}} P + & \\ \tfrac{1}{2}\sigma_{\mathrm{d}}^2\partial_{r_{\mathrm{d}}}^2 P + \tfrac{1}{2}\sigma_{\mathrm{u}}^2\partial_{r_{\mathrm{u}}}^2 P + \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}\partial_{r_{\mathrm{d}}r_{\mathrm{u}}}^2 P - r_{\mathrm{d}}P &= 0, \\ P(r_{\mathrm{d}}, r_{\mathrm{u}}, 0) &= 1. \end{aligned} \tag{1.17}$$

They expected a solution of the form

$$P(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau) = \exp\{A(\tau) - D(\tau)r_{\mathrm{d}} - U(\tau)r_{\mathrm{u}}\}. \tag{1.18}$$

This form of a solution gives $\partial_\tau P/P = \dot{A} - \dot{D}r_{\mathrm{d}} - \dot{U}r_{\mathrm{u}}$, $\partial_{r_{\mathrm{d}}} P/P = -D$, $\partial_{r_{\mathrm{d}}}^2 P/P = D^2$, $\partial_{r_{\mathrm{u}}} P/P = -U$, $\partial_{r_{\mathrm{u}}}^2 P/P = U^2$ and $\partial_{r_{\mathrm{d}}r_{\mathrm{u}}}^2 P/P = DU$. Therefore, by setting such a solution (1.18) into the PDE (1.17) we obtain that

$$\dot{D} = 1 - bD, \tag{1.19}$$

$$\dot{U} = bD - cU, \tag{1.20}$$

$$\dot{A} = (-a + \lambda_{\mathrm{d}}\sigma_{\mathrm{d}})D + (-cd + \lambda_{\mathrm{u}}\sigma_{\mathrm{u}})U + \frac{1}{2}\sigma_{\mathrm{d}}^2 D^2 + \frac{1}{2}\sigma_{\mathrm{u}}^2 U^2 + \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}DU, \tag{1.21}$$

$$A(0) = D(0) = U(0) = 0.$$

In the following we solve the previous system of ODEs, because the solution in the original paper is incorrect. The first two equations are non-homogeneous and linear. In general, let us assume the following differential equation:

$$\dot{x}(t) = \alpha x(t) + \beta(t),$$

with $x(0) = 0$ given, and β is a continuous function. Its solution is given by

$$x(t) = \exp\{\alpha t\} \int_0^t \exp\{-\alpha s\}\beta(s)\mathrm{d}s. \tag{1.22}$$

Setting the corresponding parameters of the functions D and U into formula (1.22) yields

$$D(\tau) = \frac{1 - \exp\{-b\tau\}}{b}, \tag{1.23}$$

$$U(\tau) = \begin{cases} \dfrac{b}{c-b}\left(D(\tau) - \Xi(\tau)\right), & \text{if } b \neq c \\ \Xi(\tau) - \tau\exp\{-c\tau\}, & \text{if } b = c \end{cases}, \tag{1.24}$$

where

$$\Xi(\tau) = (1 - \exp\{-c\tau\})/c. \tag{1.25}$$

A solution to A can be obtained by integrating the equation (1.21), that is

$$\begin{aligned} A(\tau) &= \int_0^\tau \bigg[(-a + \lambda_\mathrm{d}\sigma_\mathrm{d})D(s) + (-cd + \lambda_\mathrm{u}\sigma_\mathrm{u})U(s) + \frac{1}{2}\sigma_\mathrm{d}^2[D(s)]^2 + \frac{1}{2}\sigma_\mathrm{u}^2[U(s)]^2 \\ &\quad + \rho\sigma_\mathrm{d}\sigma_\mathrm{u}D(s)U(s)\bigg]\mathrm{d}s. \end{aligned}$$

It is easy to show that

$$\begin{aligned} \int_0^\tau D(s)\mathrm{d}s &= \frac{\tau}{b} - \frac{D(\tau)}{b}, \\ \int_0^\tau \Xi(s)\mathrm{d}s &= \frac{\tau}{c} - \frac{\Xi(\tau)}{c}, \\ \int_0^\tau [D(s)]^2\mathrm{d}s &= \frac{\tau}{b^2} - \frac{2}{b^2}D(\tau) - \frac{[1 - bD(\tau)]^2 - 1}{2b^3}, \\ \int_0^\tau [\Xi(s)]^2\mathrm{d}s &= \frac{\tau}{c^2} - \frac{2}{c^2}\Xi(\tau) - \frac{[1 - c\Xi(\tau)]^2 - 1}{2c^3}. \end{aligned}$$

Consequently,

$$\int_0^\tau U(s)\mathrm{d}s = \begin{cases} \dfrac{b}{c-b}\left(\dfrac{\tau}{b} - \dfrac{D(\tau)}{b} - \dfrac{\tau}{c} + \dfrac{\Xi(\tau)}{c}\right), & \text{if } b \neq c \\ \dfrac{2}{c}[\tau - \Xi(\tau)] - \tau\Xi(\tau), & \text{if } b = c \end{cases},$$

$$\int_0^\tau [U(s)]^2\mathrm{d}s = \begin{cases} \left(\dfrac{b}{c-b}\right)^2 \bigg[\dfrac{\tau}{b^2} - \dfrac{2}{b^2}D(\tau) - \dfrac{[1 - bD(\tau)]^2 - 1}{2b^3} & \text{if } b \neq c \\ \quad + \dfrac{\tau}{c^2} - \dfrac{2}{c^2}\Xi(\tau) - \dfrac{[1 - c\Xi(\tau)]^2 - 1}{2c^3} - \dfrac{2}{bc}\tau & \\ \quad + \dfrac{2}{bc}\left(D(\tau) + \Xi(\tau) + \dfrac{\exp\{-(b+c)\tau\} - 1}{b+c}\right)\bigg] & \\ \dfrac{\exp\{-2c\tau\}}{4c^3}\big[-5 + 8\exp\{c\tau\}(2 + c\tau) - 2c\tau(3 + c\tau)\big], & \text{if } b = c \\ \quad + \dfrac{4c\tau - 11}{4c^3} & \end{cases},$$

and

$$\int_0^\tau D(s)U(s)\mathrm{d}s = \begin{cases} \dfrac{2c-b}{bc(b-c)}D(\tau) - \dfrac{[2 - bD(\tau)]D(\tau)}{2b(b-c)}, & \text{if } b \neq c \\ \quad + \dfrac{1 - \exp\{-(b+c)\tau\}}{c(b^2 - c^2)} - \dfrac{\Xi(\tau)}{c(b-c)} + \dfrac{\tau}{bc} & \\ \dfrac{\exp\{-2c\tau\}}{4c^3}\big[-3 - 2c\tau + 4\exp\{c\tau\}(3 + c\tau)\big], & \text{if } b = c \\ \quad + \dfrac{4c\tau - 9}{4c^3} & \end{cases}.$$

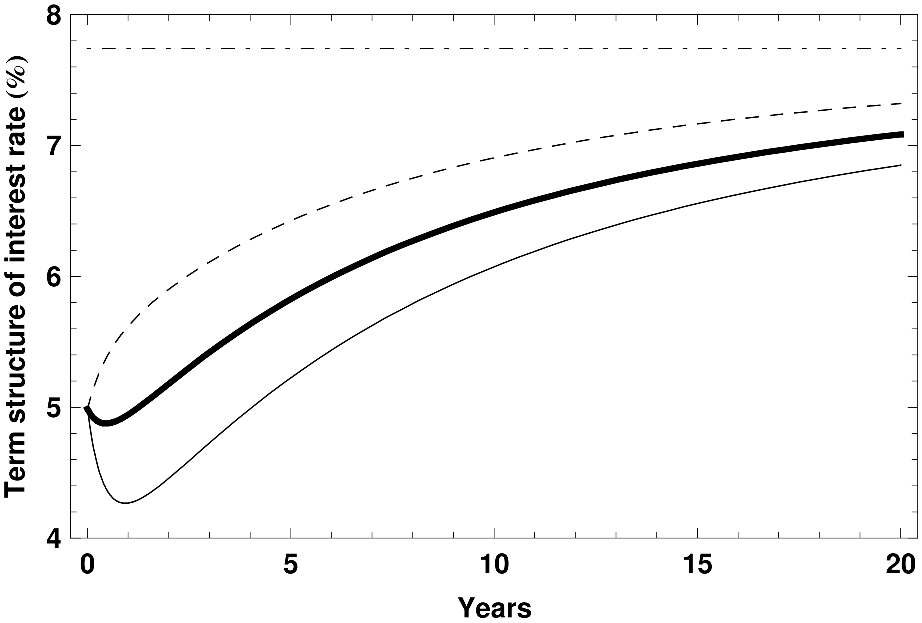

Figure 1.6: *Term structures in the two-factor convergence model of Vasicek type with parameters $a = 0.0938$, $b = 3.67$, $c = 0.2087$, $d = 0.035$, $\sigma_\mathrm{d} = 0.032$, $\sigma_\mathrm{u} = 0.016$, $\lambda_\mathrm{d} = 3.315$, $\lambda_\mathrm{u} = -0.655$, $\rho = 0.5$, $r_\mathrm{d} = 0.05$ and for r_u equal to 0.04 (solid thin line), 0.05 (solid bold line) and 0.06 (dashed thin line). The limit value of the term structure is marked with a dot-dashed line.*

In comment 8 at page 247 Corzo and Schwartz (2000) stated that "all the results can be extended in a straightforward manner to the CIR model." In Chapter 2 we show that a separable solution of the form (1.18) can only be obtained in the case of a zero correlation.

Figure 1.6 illustrates a few term structures of the domestic rate in the model by Corzo and Schwartz with parameters $a = 0.0938$, $b = 3.67$, $c = 0.2087$, $d = 0.035$, $\sigma_\mathrm{d} = 0.032$, $\sigma_\mathrm{u} = 0.016$, $\lambda_\mathrm{d} = 3.315$, $\lambda_\mathrm{u} = -0.655$, $\rho = 0.5$,$r_\mathrm{d} = 0.05$ and different values of r_u.

In the following proposition we give the limit of the domestic term structure of interest rates.

Proposition 5. *The limit of the domestic term structure of interest rates in the convergence model of Vasicek type is*

$$\lim_{\tau\to\infty} R(r_\mathrm{d}, r_\mathrm{u}, \tau) = \frac{a}{b} + d - \frac{c^2\sigma_\mathrm{d} + b^2\sigma_\mathrm{u}(2c\lambda_\mathrm{u} + \sigma_\mathrm{u}) + 2bc\sigma_\mathrm{d}(c\lambda_\mathrm{d} + \rho\sigma_\mathrm{u})}{2b^2c^2}. \tag{1.26}$$

Proof. Clearly,

$$\lim_{\tau\to\infty} D(\tau)/\tau = 0, \text{ and } \lim_{\tau\to\infty} U(\tau)/\tau = 0,$$

which follows directly from (1.23) and (1.24). Therefore, the limit of the term structure of interest rates $R(r_\mathrm{d}, r_\mathrm{u}, \tau)$ (cf. Section 1.1), after long, but straightforward computations, is

$$\begin{aligned}\lim_{\tau\to\infty} R(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau) &= \lim_{\tau\to\infty} -\frac{A(\tau)}{\tau} \\ &= \frac{a}{b} + d - \frac{c^2\sigma_{\mathrm{d}} + b^2\sigma_{\mathrm{u}}(2c\lambda_{\mathrm{u}} + \sigma_{\mathrm{u}}) + 2bc\sigma_{\mathrm{d}}(c\lambda_{\mathrm{d}} + \rho\sigma_{\mathrm{u}})}{2b^2c^2}.\end{aligned}$$

□

The previous proposition tells us that the limit of the term structures in Figure 1.6 is $R_\infty = 0.0774$. Note that the limit of the term structure does not depend on r_{d} and r_{u} (similarly to the one-factor model).

2

Two-factor convergence model of Cox-Ingersoll-Ross type

2.1 Motivation for the model and its formulation

In the previous chapter we discussed a few well-known two-factor models. In particular, we focused our attention on the convergence model proposed by Corzo and Schwartz (2000). The authors applied the Vasicek model to obtain a convergence model for a country before adopting the Euro currency. A disadvantage of the two-factor convergence model of Vasicek type is that it allows negative values of the domestic and EMU's interest rates.

In this chapter we analyse the *two-factor convergence model of CIR type*. More precisely, the EMU's short-rate r_{u} and the domestic short-rate r_{d} are assumed to be linked in the following way:

$$\begin{aligned} \mathrm{d}r_{\mathrm{d}} &= [a + b(r_{\mathrm{u}} - r_{\mathrm{d}})]\mathrm{d}t + \sigma_{\mathrm{d}}\sqrt{r_{\mathrm{d}}}\mathrm{d}W_{\mathrm{d}}, \\ \mathrm{d}r_{\mathrm{u}} &= c(d - r_{\mathrm{u}})\mathrm{d}t + \sigma_{\mathrm{u}}\sqrt{r_{\mathrm{u}}}\mathrm{d}W_{\mathrm{u}}, \\ \mathrm{Cov}[\mathrm{d}W_{\mathrm{d}}, \mathrm{d}W_{\mathrm{u}}] &= \rho \mathrm{d}t, \end{aligned} \tag{2.1}$$

where $\mathrm{d}W_{\mathrm{d}}$ and $\mathrm{d}W_{\mathrm{u}}$ are increments of Wiener processes, or, equivalently,

$$\begin{aligned} \mathrm{d}r_{\mathrm{d}} &= [a + b(r_{\mathrm{u}} - r_{\mathrm{d}})]\mathrm{d}t + \sigma_{\mathrm{d}}\sqrt{r_{\mathrm{d}}}\left[\sqrt{1-\rho^2}\mathrm{d}\widetilde{W}_{\mathrm{d}} + \rho\mathrm{d}\widetilde{W}_{\mathrm{u}}\right], \\ \mathrm{d}r_{\mathrm{u}} &= c(d - r_{\mathrm{u}})\mathrm{d}t + \sigma_{\mathrm{u}}\sqrt{r_{\mathrm{u}}}\mathrm{d}\widetilde{W}_{\mathrm{u}}, \end{aligned} \tag{2.2}$$

where $\mathrm{d}\widetilde{W}_{\mathrm{d}}$ and $\mathrm{d}\widetilde{W}_{\mathrm{u}}$ are increments of independent Wiener processes. The process for the EMU's rate r_{u} is a mean-reversion process with a limit $d > 0$. The process for the domestic rate r_{d} converges to r_{u} with a possible minor divergence given by

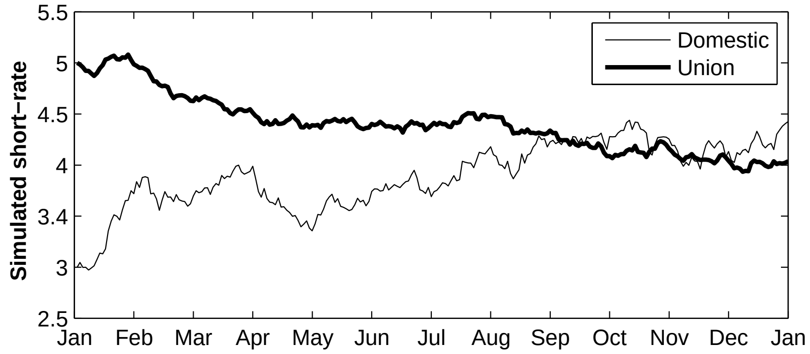

Figure 2.1: *A simulation of the process* (2.1) *with parameters* $a = 0.01$, $b = 3.67$, $c = 0.21$, $d = 0.03$, $\sigma_{\mathrm{d}} = 0.05$, $\sigma_{\mathrm{u}} = 0.02$ *and* $\rho = 0.22$. *The initial values are 3 percent for the domestic and 5 percent for the EMU's rate.*

a. The coefficients $b > 0$ and $c > 0$ describe the speed of the convergence. The volatilities are determined by positive constants σ_{d} and σ_{u} multiplied by the square root of the corresponding value of the interest rate. A simulation of possible paths of the domestic and EMU's short-rates is illustrated in Figure 2.1. Compared with the convergence model of Vasicek type, the model (2.1) rejects the possibility of a negative value of the interest rates.

2.2 A solution to the bond-pricing PDE

Let the EMU's market price of risk be equal to $\nu_{\mathrm{u}}\sqrt{r_{\mathrm{u}}}$, where ν_{u} is a constant. Then the price of the EMU's discount bond is given by the CIR bond-pricing formula (see, Cox et al. (1985), or Section 1.3). By setting the corresponding drifts and volatilities to the bond-pricing PDE (1.16), we obtain that the price of the domestic bond $P(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau)$ is a solution to

$$\begin{aligned} -\partial_\tau P + [a + b(r_{\mathrm{u}} - r_{\mathrm{d}}) - \lambda_{\mathrm{d}}(r_{\mathrm{d}}, r_{\mathrm{u}})\sigma_{\mathrm{d}}\sqrt{r_{\mathrm{d}}}]\partial_{r_{\mathrm{d}}} P + [c(d - r_{\mathrm{u}}) - \nu_{\mathrm{u}}\sigma_{\mathrm{u}} r_{\mathrm{u}}]\partial_{r_{\mathrm{u}}} P \\ + \frac{\sigma_{\mathrm{d}}^2 r_{\mathrm{d}}}{2} \cdot \partial^2_{r_{\mathrm{d}}} P + \frac{\sigma_{\mathrm{u}}^2 r_{\mathrm{u}}}{2} \cdot \partial^2_{r_{\mathrm{u}}} P + \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}\sqrt{r_{\mathrm{d}} r_{\mathrm{u}}}\partial^2_{r_{\mathrm{d}} r_{\mathrm{u}}} P - r_{\mathrm{d}} P = 0, \end{aligned} \quad (2.3)$$

where $\lambda_{\mathrm{d}}(r_{\mathrm{d}}, r_{\mathrm{u}})$ is the domestic market price of risk.

The case of a zero correlation

Let be the domestic market price of risk taken to be $\lambda_{\mathrm{d}}(r_{\mathrm{d}}, r_{\mathrm{u}}) = \nu_{\mathrm{d}}\sqrt{r_{\mathrm{d}}}$, where ν_{d} is a constant (i.e., the domestic market price of risk has the same functional form as the EMU's one). Then PDE (2.3) yields that the price of a discount bond solves

$$\begin{aligned} -\partial_\tau P + [a + b(r_{\mathrm{u}} - r_{\mathrm{d}}) - \nu_{\mathrm{d}}\sigma_{\mathrm{d}} r_{\mathrm{d}}]\partial_{r_{\mathrm{d}}} P + [c(d - r_{\mathrm{u}}) - \nu_{\mathrm{u}}\sigma_{\mathrm{u}} r_{\mathrm{u}}]\partial_{r_{\mathrm{u}}} P \\ + \frac{\sigma_{\mathrm{d}}^2 r_{\mathrm{d}}}{2}\partial^2_{r_{\mathrm{d}}} P + \frac{\sigma_{\mathrm{u}}^2 r_{\mathrm{u}}}{2}\partial^2_{r_{\mathrm{u}}} P - r_{\mathrm{d}} P \ = \ 0. \end{aligned} \quad (2.4)$$

A solution to the bond-pricing PDE in the case of zero correlation

Let us assume that the solution of the previous PDE has the form of (1.18), that is

$$P(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau) = \exp\{A(\tau) - D(\tau)r_{\mathrm{d}} - U(\tau)r_{\mathrm{u}}\}.$$

We repeat that in this case $P > 0$ and $\partial_\tau P/P = \dot{A} - \dot{D}r_{\mathrm{d}} - \dot{U}r_{\mathrm{u}}$, $\partial_{r_{\mathrm{d}}} P/P = -D$, $\partial^2_{r_{\mathrm{d}}} P/P = D^2$, $\partial_{r_{\mathrm{u}}} P/P = -U$, $\partial^2_{r_{\mathrm{u}}} P/P = U^2$ and $\partial^2_{r_{\mathrm{d}} r_{\mathrm{u}}} P/P = DU$. Setting such a solution to PDE (2.4) gives us that

$$\begin{aligned} -(\dot{A} - \dot{D}r_{\mathrm{d}} - \dot{U}r_{\mathrm{u}}) + [a + b(r_{\mathrm{u}} - r_{\mathrm{d}}) - \nu_{\mathrm{d}}\sigma_{\mathrm{d}}r_{\mathrm{d}}](-D) + [c(d - r_{\mathrm{u}}) - \nu_{\mathrm{u}}\sigma_{\mathrm{u}}r_{\mathrm{u}}](-U) \\ + \frac{\sigma_{\mathrm{d}}^2 r_{\mathrm{d}}}{2} D^2 + \frac{\sigma_{\mathrm{u}}^2 r_{\mathrm{u}}}{2} U^2 - r_{\mathrm{d}} P = 0, \end{aligned}$$

and after some transformation we obtain

$$\begin{aligned} (-\dot{A} - aD - cdU) + r_{\mathrm{d}} \left[\dot{D} - 1 + (b + \nu_{\mathrm{d}}\sigma_{\mathrm{d}})D + \tfrac{\sigma_{\mathrm{d}}^2}{2} D^2\right] + \\ r_{\mathrm{u}} \left[\dot{U} - bD + (c + \nu_{\mathrm{u}}\sigma_{\mathrm{u}})U + \tfrac{\sigma_{\mathrm{u}}^2}{2} U^2\right] &= 0. \end{aligned}$$

Since the previous equation holds for any r_{d}, r_{u}, functions A, D and U solve the system of ODEs

$$\dot{D} = 1 - (b + \nu_{\mathrm{d}}\sigma_{\mathrm{d}})D - \frac{\sigma_{\mathrm{d}}^2}{2} D^2, \tag{2.5}$$

$$\dot{U} = bD - (c + \nu_{\mathrm{u}}\sigma_{\mathrm{u}})U - \frac{\sigma_{\mathrm{u}}^2}{2} U^2, \tag{2.6}$$

$$\dot{A} = -aD - cdU, \tag{2.7}$$

with initial conditions $A(0) = D(0) = U(0) = 0$.

First, we solve ODE (2.5). By ϕ we denote the term $\sigma_{\mathrm{d}}^2/2$, and by ψ we denote the term $b + \nu_{\mathrm{d}}\sigma_{\mathrm{d}}$. Clearly, a solution to ODE (2.5) follows from

$$\int_0^\tau \frac{\mathrm{d}D}{1 - \psi D - \phi D^2} = \tau - C, \tag{2.8}$$

where C is a constant. Since both ϕ and ψ are positive, then $\psi^2 + 4\phi > 0$, and the denominator of the left-hand term has two roots $D_\oplus$ and $D_\ominus$:

$$D_{\ominus,\oplus} = \frac{\psi \pm \sqrt{\psi^2 + 4\phi}}{-2\phi}. \tag{2.9}$$

It is easy to see that $D_\oplus < 0$ and $D_\ominus > 0$. We decompose the left-hand fraction into the sum of two fractions, i.e.

$$\frac{1}{1 - \psi D - \phi D^2} = \frac{(Q_\oplus + Q_\ominus)D - Q_\oplus D_\ominus - Q_\ominus D_\oplus}{-\phi(D - D_\oplus)(D - D_\ominus)}.$$

The previous equality yields that $Q_\ominus = -Q_\oplus$, which implies that $Q_\oplus = (D_\oplus - D_\ominus)^{-1} = -\phi(\psi^2 + 4\phi)^{-1/2}$. Evidently, $k = -\phi/Q_\oplus = \sqrt{\psi^2 + 4\phi}$. Hence

$$\frac{1}{1 - \psi D - \phi D^2} = \frac{1}{k}\left(\frac{1}{D - D_\oplus} - \frac{1}{D - D_\ominus}\right).$$

By applying the previous equation to solution (2.8), we obtain that a solution to the ODE (2.5) satisfies

$$\frac{1}{k}\ln\left|\frac{D-D_\oplus}{D-D_\ominus}\right| + C = \tau$$

The initial condition $D(0) = 0$ and the fact that the fraction in the logarithm has a negative value enable us to express an explicit solution to ODE (2.5)

$$D(\tau) = \frac{D_\oplus(1-\exp\{k\tau\})}{1-\frac{D_\oplus}{D_\ominus}\exp\{k\tau\}}. \tag{2.10}$$

We were not able to find an explicit solution to ODE (2.6); nevertheless, it is easy to solve it numerically and obtain the values of $U(\tau)$. The function $A(\tau)$ can then be obtained by a numerical integration of equation (2.7).

Properties of a solution in the case of zero correlation

The following statements formulate some properties of functions A, D and U.

Lemma 6. *Let $A(\tau)$, $D(\tau)$ and $U(\tau)$ be solutions to the system of ODEs* (2.5)–(2.7). *Then:* ***i)*** *$D(\tau) > 0$ is monotonous and increasing, and $\lim_{\tau\to\infty} D(\tau) = D_\ominus$,* ***ii)*** *$U(\tau) > 0$ is monotonous, increasing and bounded, and* ***iii)*** *if $a \geq 0$ then $A(\tau) < 0$; for all $\tau > 0$.*

Proof. **i)** The monotonicity of D follows directly from the derivative of solution (2.10) with respect to τ, which is positive:

$$\dot{D}(\tau) = \frac{-D_\oplus(1-\frac{D_\oplus}{D_\ominus})k\exp\{k\tau\}}{(1-\frac{D_\oplus}{D_\ominus}\exp\{k\tau\})^2} > 0,$$

since $k > 0$, $D_\oplus < 0$, and $D_\ominus > 0$. The fact that $D(0) = 0$ and $\dot{D}(\tau) > 0$ for τ greater than 0 implies the positivity of D. For $\tau \to \infty$ we obtain:

$$\begin{aligned}\lim_{\tau\to\infty} D(\tau) &= \lim_{\tau\to\infty}\left[\frac{D_\oplus}{1-\frac{D_\oplus}{D_\ominus}\exp\{k\tau\}} - \frac{D_\oplus\exp\{k\tau\}}{1-\frac{D_\oplus}{D_\ominus}\exp\{k\tau\}}\right] \\ &= \lim_{\tau\to\infty}\left[\frac{-D_\oplus}{\exp\{-k\tau\}-D_\oplus/D_\ominus}\right] = D_\ominus.\end{aligned}$$

ii) The initial condition $U(0) = 0$ and equation (2.6) imply that $\dot{U}(0) = 0$ and $\ddot{U}(0) = b > 0$. Therefore, U is positive in some neighbourhood of $\tau = 0$. To prove the positivity of U for all τ greater than 0, it is sufficient to show that $\dot{U}(\tau^*) > 0$ whenever $U(\tau^*) = 0$. This holds since if $U(\tau^*) = 0$, then, due to equation (2.6), we obtain $\dot{U}(\tau^*) = bD(\tau^*) > 0$. To prove that U is monotonous and increasing, we have to show that $\dot{U}$ is positive. To do this we show that if $\dot{U}(\tau^*) = 0$, then $\ddot{U}(\tau^*) = b\dot{D}(\tau^*) - (c+\nu_\mathrm{u}\sigma_\mathrm{u})\dot{U}(\tau^*) - \frac{\sigma_\mathrm{u}^2}{2}U(\tau^*)\dot{U}(\tau^*) = b\dot{D}(\tau^*) > 0$. To prove that U is bounded it is sufficient to show that there exists M such that if $U(\tau^*) = M > 0$, then $\dot{U}(\tau^*) \leq 0$: $\dot{U}(\tau^*) = bD(\tau^*) - (c+\nu_\mathrm{u}\sigma_\mathrm{u})M - \frac{\sigma_\mathrm{u}^2}{2}M^2 \leq 0$. **iii)** Since $D(\tau) > 0$ and $U(\tau) > 0$ for all $\tau > 0$, equation (2.7) implies that $\dot{A}(\tau) < 0$ for all $\tau > 0$, i.e., $A(\tau)$ is strictly decreasing with an origin in 0, which proves the third part. □

The previous lemma implies

Corollary 7. *The limit of $U(\tau)$ for $\tau \to \infty$ is*

$$\widehat{U} = \lim_{\tau\to\infty} U(\tau) = \frac{(c+\sigma_{\mathrm{u}}\nu_{\mathrm{u}}) - \sqrt{(c+\sigma_{\mathrm{u}}\nu_{\mathrm{u}})^2 + 2b\sigma_{\mathrm{u}}^2 D_{\ominus}}}{-\sigma_{\mathrm{u}}^2}.$$

Proof. The boundedness and monotonicity of U gives the existence of $\lim_{\tau\to\infty} U(\tau)$, and, consequently, $\lim_{\tau\to\infty} \dot{U}(\tau)$ (cf. equation (2.6)). Let $\lim_{\tau\to\infty} \dot{U}(\tau) = L \neq 0$. Then, from the definition of a limit, there exists K such that for all $\tau \in (K, \infty)$ we have $\dot{U}(\tau) \geq L/2$. Langrange's mean value theorem yields that for any $s, t \in (K, \infty)$, we have $U(s) - U(t) = \dot{U}(\tau)(s-t) \geq \frac{L}{2}(s-t)$. Therefore, $U(s) \geq U(t) + \frac{L}{2}(s-t)$, which implies a contradiction for $s \to \infty$. Consequently, $\lim_{\tau\to\infty} \dot{U}(\tau) = 0$. We obtain that

$$\lim_{\tau\to\infty} \dot{U}(\tau) = 0 = bD_{\ominus} - (c+\sigma_{\mathrm{u}}\nu_{\mathrm{u}})\widehat{U} - \frac{\sigma_{\mathrm{u}}^2}{2}\widehat{U}^2.$$

The positive solution of the previous equation is the limit of U. □

It follows that if $a \geq 0$, then the bond price lies in the interval $(0,1)$ for all $\tau > 0$; hence, the term structures starting from the positive short-rate are always positive. Note that this is not necessarily true in two-factor models; Stehlíková and Ševčovič (2005) showed that a certain constraint on the market price of risk has to be imposed to ensure the positivity of the interest rates in the Fong-Vasicek model.

Figure 2.2 illustrates a few examples of term structures obtained from the convergence model of CIR type with zero correlation.

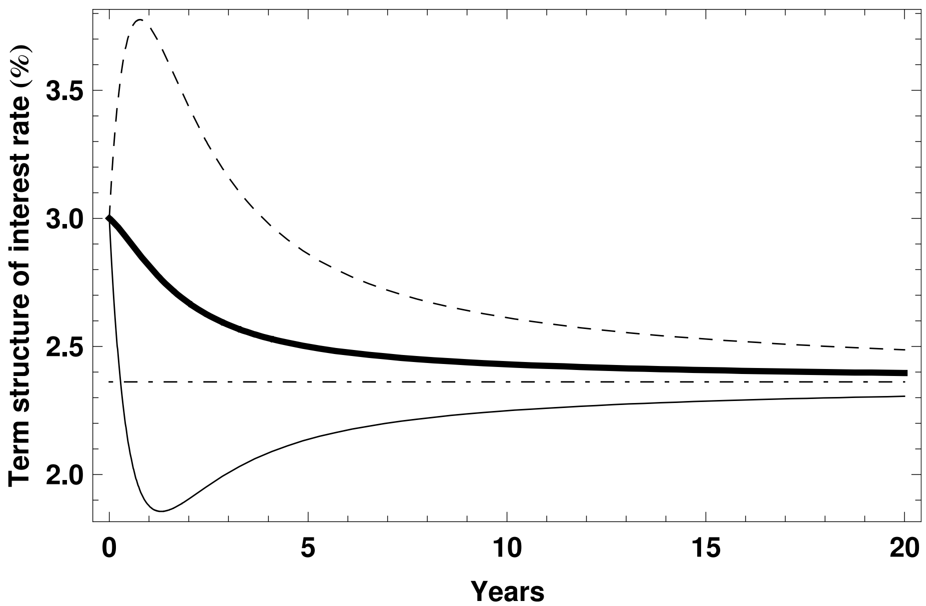

Figure 2.2: *Term structures in the two-factor convergence model of CIR with parameters $a = -0.01$, $b = 3$, $c = 1$, $d = 0.03$, $\sigma_{\mathrm{d}} = 0.05$, $\sigma_{\mathrm{u}} = 0.04$, $\nu_{\mathrm{d}} = -5$, $\nu_{\mathrm{u}} = 5$, $r_{\mathrm{d}} = 0.03$ and for r_{u} equal to 0.01 (solid thin line), 0.03 (solid bold line) and 0.05 (dashed thin line). The limit value of the term structure is marked with a dot-dashed line.*

In the following proposition we state the limit of the domestic term structure of interest rates in the convergence model of CIR type.

Proposition 8. *The limit of the domestic term structure of interest rates in the convergence model of CIR type is*

$$\lim_{\tau\to\infty} R(r_{\rm d}, r_{\rm u}, \tau) = aD_{\ominus} + cd\frac{(c+\nu_{\rm u}\sigma_{\rm u}) - \sqrt{(c+\nu_{\rm u}\sigma_{\rm u})^2 + 2bD_{\ominus}\sigma_{\rm u}^2}}{-\sigma_{\rm u}^2},$$

where $D_{\ominus}$ *is defined by* (2.9).

Proof. Lemma 6 and Corollary 7 imply that

$$\lim_{\tau\to\infty} \frac{D(\tau)}{\tau} = 0, \text{ and } \lim_{\tau\to\infty} \frac{U(\tau)}{\tau} = 0.$$

Using l'Hospital's rule and ODE (2.7), the limit of the term structure is

$$\lim_{\tau\to\infty} R(r_{\rm d}, r_{\rm u}, \tau) = -\lim_{\tau\to\infty} A(\tau)/\tau = -\lim_{\tau\to\infty} \dot{A}(\tau) = a\lim_{\tau\to\infty} D(\tau) + cd\lim_{\tau\to\infty} U(\tau),$$

which completes the proof. □

According to the previous proposition, the limit value of the term structures in Figure 2.2 is $R_\infty = 0.0236$.

The case of a nonzero correlation

In the case of a nonzero correlation the term $\rho\sigma_{\rm d}\sigma_{\rm u}\sqrt{r_{\rm d}r_{\rm u}}\partial^2_{r_{\rm d}r_{\rm u}}P$ in equation (2.3) is not eliminated. The only acceptable domestic market price of risk is of the form $\lambda_{\rm d} = \nu_{\rm d}\sqrt{r_{\rm d}} + \overline{\nu_{\rm u}}\sqrt{r_{\rm u}}$, where $\nu_{\rm d}$ and $\overline{\nu_{\rm u}}$ are constants (this approach enables us to obtain one more term with $\sqrt{r_{\rm d}r_{\rm u}}$; the other choice would lead to a single term that we would not be able to eliminate). If we assume the solution of the form (1.18), the only change is that the system of ODEs (2.5)–(2.7) is extended by the equation

$$0 = \overline{\nu_{\rm u}}\sigma_{\rm d}D + \rho\sigma_{\rm d}\sigma_{\rm u}DU. \tag{2.11}$$

However, Lemma 6 implies that in the solution of the form (1.18) the function D is positive. It is obvious that U is not a constant function; therefore, equation (2.11) is not satisfied for $\rho \neq 0$.

2.3 Approximation and accuracy

Although we proved that there is no separable solution of the form (1.18), we can try to approximate a solution in the nonzero correlation case by a solution in the zero correlation case. In this section we also investigate how much these solutions differ. The following example demonstrates our motivation.

Maturity	Difference
1/4	4.5×10^{-5}
1/2	2.2×10^{-4}
3/4	4.8×10^{-4}
1	7.8×10^{-4}
5	5.0×10^{-3}
10	8.1×10^{-3}
20	1.1×10^{-2}

Table 2.1: *Differences (in percent) in interest rates between the convergence model of Vasicek type with and without a correlation. The parameters of the model were taken from Corzo and Schwartz (2000).*

Motivation: the convergence model of Vasicek type

Let us consider the two-factor convergence model of Vasicek type (see Section 1.4); for the sake of simplicity assume that $c \neq d$. Let $P_{\text{Vas}}(r_{\text{d}}, r_{\text{u}}, \tau; \rho)$ be the price of the domestic bond, where the dependence on the correlation ρ is explicitly marked (analogously, let $R(r_{\text{d}}, r_{\text{u}}, \tau; \rho)$ be the term structure of interest rates and $A(\tau; \rho)$ be the function in (1.18)). By expanding the explicit solution into the Taylor series with respect to τ we obtain that:

Proposition 9. *Let $P_{\text{Vas}}(r_{\text{d}}, r_{\text{u}}, \tau; \rho)$ be a solution to the bond-pricing PDE* (1.17) *of the convergence model of Vasicek type. Then*

$$\ln[P_{\text{Vas}}(r_{\text{d}}, r_{\text{u}}, \tau; 0)] - \ln[P_{\text{Vas}}(r_{\text{d}}, r_{\text{u}}, \tau; \rho)] = -\frac{1}{8} b\rho\sigma_{\text{d}}\sigma_{\text{u}}\, \tau^4 + \text{o}(\tau^4).$$

In Table 2.1 we exhibit the difference between interest rates in the convergence model of Vasicek type with parameters taken from Corzo and Schwartz (2000) with the same model with a zero correlation and the other parameters remaining. The market data are quoted with two decimal places; therefore, the differences in Table 2.1 are observable only for long-time maturities. However, even in the case of a twenty-year maturity, the difference is only 0.01 percent (for the given parameters).

The question that arises is: what is the maximal possible difference in the convergence model of Vasicek type. The result is stated in the following

Lemma 10. *In the convergence model of Vasicek type, the difference $|R(r_{\text{d}}, r_{\text{u}}, \tau; 0) - R(r_{\text{d}}, r_{\text{u}}, \tau; \rho)|$ between the term structures of interest rates in the case of the zero and nonzero correlation is monotonous and less than or equal to $|\rho|\sigma_{\text{d}}\sigma_{\text{u}}/(bc)$.*

Proof. Let us denote the difference between the term structures of interest rates with and without a correlation (cf. Section 1.1)

$$R(\tau) = R(r_{\text{d}}, r_{\text{u}}, \tau; 0) - R(r_{\text{d}}, r_{\text{u}}, \tau; \rho) = -\frac{A(\tau; 0) - A(\tau; \rho)}{\tau}.$$

First we show that $R(\tau)$ is monotonous. By differentiating $R(\tau)$ and applying ODE (1.21) we get

$$\dot{R}(\tau) = -\frac{[\dot{A}(\tau;0) - \dot{A}(\tau;\rho)]\tau - [A(\tau;0) - A(\tau;\rho)]}{\tau^2} = -\frac{1}{\tau}R + \frac{\rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}D(\tau)U(\tau)}{\tau}. \tag{2.12}$$

The variation of constants method yields that the solution of the ODE (2.12) is $R(\tau) = R_{\mathrm{H}}(\tau)c(\tau)$, where $R_{\mathrm{H}}(\tau) = 1/\tau$ and $\dot{c}(\tau) = \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}D(\tau)U(\tau)$. Consequently (the integrating constant has to be equal to zero),

$$R(\tau) = \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}F(\tau),$$

where

$$F(\tau) = \frac{1}{\tau}\int_0^\tau D(s)U(s)\mathrm{d}s.$$

To prove that R is monotonous it is sufficient to show that $\dot{F}$ is positive for $\tau > 0$. The derivative of F is

$$\dot{F}(\tau) = \frac{D(\tau)U(\tau)\tau - \int_0^\tau D(s)U(s)\mathrm{d}s}{\tau^2}.$$

Since D and U are both increasing and positive, then the product DU is increasing and positive, too. Therefore,

$$\int_0^\tau D(s)U(s)\mathrm{d}s < \int_0^\tau \max_{t\in\langle 0,\tau\rangle} D(t)U(t)\mathrm{d}s = \int_0^\tau D(\tau)U(\tau)\mathrm{d}s = D(\tau)U(\tau)\tau,$$

which implies that $\dot{F} > 0$ for τ greater than 0. The maximum possible value of $|R|$ is

$$\max_{\tau\in(0,\infty)} |R(\tau)| \;=\; |\rho|\sigma_{\mathrm{d}}\sigma_{\mathrm{u}} \max_{\tau\in(0,\infty)} F(\tau) = |\rho|\sigma_{\mathrm{d}}\sigma_{\mathrm{u}} \lim_{\tau\to\infty} F(\tau) = |\rho|\sigma_{\mathrm{d}}\sigma_{\mathrm{u}} \lim_{\tau\to\infty} D(\tau)U(\tau).$$

To complete the proof we set the limits $\lim_{\tau\to\infty} D(\tau) = 1/b$ (cf. equation (1.23)), and $\lim_{\tau\to\infty} U(\tau) = 1/c$ (cf. equation (1.24)). □

This approach motivates us to determine the difference between the logarithm of the price of the discount bond in the case of a zero correlation and the logarithm of the price of the discount bond in the case of a nonzero correlation in the two-factor convergence model of CIR type.

Approximation of a solution in the case of a nonzero correlation in the convergence model of CIR type and its accuracy

Since we do not know any exact solution to the discount bond price in the convergence model of CIR type, we are only able to derive the order of an approximation. The result is formulated in the following

Theorem 11. *Let $P_{\mathrm{CIR}}(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau; \rho)$ be a solution to the bond-pricing PDE* (2.3) *of the convergence model of CIR type. Then*

$$\ln[P_{\mathrm{CIR}}(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau; 0)] - \ln[P_{\mathrm{CIR}}(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau; \rho)] = c_3(r_{\mathrm{d}}, r_{\mathrm{u}}; \rho)\tau^3 + \mathrm{o}(\tau^3),$$

where the coefficient c_3 is not identically equal to zero.

Proof. Let $f = \ln(P)$ be the logarithm of the domestic bond-price, and let $K_{\mathrm{d}} = [a + b(r_{\mathrm{u}} - r_{\mathrm{d}}) - \nu_{\mathrm{d}}\sigma_{\mathrm{d}}r_{\mathrm{d}}]$, $L_{\mathrm{d}} = \sigma_{\mathrm{d}}^2 r_{\mathrm{d}}/2$, $K_{\mathrm{u}} = [c(d - r_{\mathrm{u}}) - \nu_{\mathrm{u}}\sigma_{\mathrm{u}}r_{\mathrm{u}}]$, $L_{\mathrm{u}} = \sigma_{\mathrm{u}}^2 r_{\mathrm{u}}/2$. Then f satisfies the following PDE:

$$\begin{aligned} -\partial_\tau f + K_{\mathrm{d}}\partial_{r_{\mathrm{d}}} f + K_{\mathrm{u}}\partial_{r_{\mathrm{u}}} f + L_{\mathrm{d}}\left[(\partial_{r_{\mathrm{d}}} f)^2 + \partial^2_{r_{\mathrm{d}}} f\right] + L_{\mathrm{u}}\left[(\partial_{r_{\mathrm{u}}} f)^2 + \partial^2_{r_{\mathrm{u}}} f\right] & \\ +2\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}\left(\partial_{r_{\mathrm{d}}} f \partial_{r_{\mathrm{u}}} f + \partial^2_{r_{\mathrm{d}}r_{\mathrm{u}}} f\right) - r_{\mathrm{d}} &= 0, \end{aligned} \tag{2.13}$$

which follows from (2.3). For our purposes, we denote by P_{ex} an exact solution to equation (2.3) for $\rho > 0$; we denote by P_{ap} our solution to equation (2.3) with $\rho = 0$ by which we want to approximate P_{ex}, and $f = \ln(P_{\mathrm{ex}})$ and $f_0 = \ln(P_{\mathrm{ap}})$. Let us see what PDE $g = f_0 - f$ satisfies. Using $(\partial_r g)^2 = (\partial_r f_0)^2 - (\partial_r f)^2 - 2\partial_r f \partial_r g$, for $r = r_{\mathrm{d}}, r_{\mathrm{u}}$, we obtain

$$\begin{aligned} &-\partial_\tau g + K_{\mathrm{d}}\partial_{r_{\mathrm{d}}} g + K_{\mathrm{u}}\partial_{r_{\mathrm{u}}} g + L_{\mathrm{d}}\left[(\partial_{r_{\mathrm{d}}} g)^2 + \partial^2_{r_{\mathrm{d}}} g\right] + L_{\mathrm{u}}\left[(\partial_{r_{\mathrm{u}}} g)^2 + \partial^2_{r_{\mathrm{u}}} g\right] \\ &+2\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}\left(\partial_{r_{\mathrm{d}}} g \partial_{r_{\mathrm{u}}} g + \partial^2_{r_{\mathrm{d}}r_{\mathrm{u}}} g\right) = \\ &-\partial_\tau f_0 + K_{\mathrm{d}}\partial_{r_{\mathrm{d}}} f_0 + K_{\mathrm{u}}\partial_{r_{\mathrm{u}}} f_0 + L_{\mathrm{d}}\left[(\partial_{r_{\mathrm{d}}} f_0)^2 + \partial^2_{r_{\mathrm{d}}} f_0\right] + L_{\mathrm{u}}\left[(\partial_{r_{\mathrm{u}}} f_0)^2 + \partial^2_{r_{\mathrm{u}}} f_0\right] - r_{\mathrm{d}} \\ &- \big(-\partial_\tau f + K_{\mathrm{d}}\partial_{r_{\mathrm{d}}} f + K_{\mathrm{u}}\partial_{r_{\mathrm{u}}} f + L_{\mathrm{d}}\left[(\partial_{r_{\mathrm{d}}} f)^2 + \partial^2_{r_{\mathrm{d}}} f\right] \\ &+L_{\mathrm{u}}\left[(\partial_{r_{\mathrm{u}}} f)^2 + \partial^2_{r_{\mathrm{u}}} f\right] + 2\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}\left(\partial_{r_{\mathrm{d}}} f \partial_{r_{\mathrm{u}}} f + \partial^2_{r_{\mathrm{d}}r_{\mathrm{u}}} f\right) - r_{\mathrm{d}}\big) \\ &-2L_{\mathrm{d}}\partial_{r_{\mathrm{d}}} f \partial_{r_{\mathrm{d}}} g - 2L_{\mathrm{u}}\partial_{r_{\mathrm{u}}} f \partial_{r_{\mathrm{u}}} g + 2\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}\left(\partial_{r_{\mathrm{d}}} f \partial_{r_{\mathrm{u}}} f + \partial^2_{r_{\mathrm{d}}r_{\mathrm{u}}} f\right) \\ &+2\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}\left(\partial_{r_{\mathrm{d}}} g \partial_{r_{\mathrm{u}}} g + \partial^2_{r_{\mathrm{d}}r_{\mathrm{u}}} g\right) \\ &= 4\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}DU + 2L_{\mathrm{d}}\left[(\partial_{r_{\mathrm{d}}} f)^2 + D\partial_{r_{\mathrm{d}}} f\right] + 2L_{\mathrm{u}}\left[(\partial_{r_{\mathrm{u}}} f)^2 + U\partial_{r_{\mathrm{u}}} f\right] \\ &+2\rho\sqrt{L_{\mathrm{d}}L_{\mathrm{u}}}\left(2\partial_{r_{\mathrm{d}}} f \partial_{r_{\mathrm{u}}} f + D\partial_{r_{\mathrm{u}}} f + U\partial_{r_{\mathrm{d}}} f\right). \end{aligned} \tag{2.14}$$

Now, we expand g into the Taylor series, i.e., $g(\tau, r_{\mathrm{d}}, r_{\mathrm{u}}) = \sum_{k=\omega}^{\infty} c_k(r_{\mathrm{d}}, r_{\mathrm{u}})\tau^k$; that is, we expect the first $\omega - 1$ terms to be zero. Therefore, $\partial_\tau g = \omega c_\omega \tau^{\omega-1} + \mathrm{o}(\tau^{\omega-1})$. The rest of the terms on the left-hand side of (2.14) are of the order τ^ω (because the rest are derivatives of g with respect to r_{d} and r_{u}); hence, the left-hand side is of the order $\tau^{\omega-1}$. Let us analyse the right-hand side of the equation (2.14). Note that f is of the order τ, since its value for $\tau = 0$ is the logarithm of the bond price at maturity, i.e., zero. It follows that the derivatives $\partial_{r_{\mathrm{d}}} f$ and $\partial f r_{\mathrm{u}}$ are of the order τ as well. Equation (2.5) and the initial condition $D(0) = 0$ give $\dot{D}(0) = 1$ and $\ddot{D}(0) = -(b + \nu_{\mathrm{d}}\sigma_{\mathrm{d}})$. Analogously, equation (2.6) and $U(0) = 0$ yield that $\dot{U}(0) = 0$ and $\ddot{U}(0) = b$. Therefore, we obtain the expansion $D(\tau)U(\tau) = \frac{1}{2}b\tau^3 + \mathrm{o}(\tau^3)$. Consequently, we get that the right-hand side of equation (2.14) is of the order at least τ^2. Therefore, ω is at least 3. An order higher than 3 would be attained if the coefficient at τ^2 in the expansion of the right-hand side of (2.14) was eliminated. In the following we show that that is not the case. Since $U = b\tau^2 + \mathrm{o}(\tau^2)$, $b > 0$, and $\partial_{r_{\mathrm{u}}} g = -U - \partial_{r_{\mathrm{u}}} f$, we have extra information that $\partial_{r_{\mathrm{u}}} f = k_2\tau^2 + \mathrm{o}(\tau^2)$. Repeating the previous analysis of the right-hand side of (2.14) with this additional information, we obtain that the only $\mathrm{O}(\tau^2)$ term is (up to a multiplicative constant independent of τ) equal to $\partial_{r_\mathrm{d}} f\,(\partial_{r_\mathrm{d}} f + D)$. We prove that this term is not constantly equal to zero. To derive a contradiction, assume that

$$\partial_{r_\mathrm{d}} f\,(\partial_{r_\mathrm{d}} f + D) = 0, \tag{2.15}$$

for all $\tau, r_\mathrm{d}, r_\mathrm{u}$. It follows from the continuity of $\partial_{r_\mathrm{d}} f$ and the behaviour of D that we only have two options: either $\partial_{r_\mathrm{d}} f = 0$ for all $\tau, r_\mathrm{d}, r_\mathrm{u}$, or $\partial_{r_\mathrm{d}} f + D = 0$ for all $\tau, r_\mathrm{d}, r_\mathrm{u}$. If $\partial_{r_\mathrm{d}} f = 0$, then $f = f(r_\mathrm{u}, \tau)$ and $\partial^2_{r_\mathrm{d}} f = 0$. Equation (2.13) reduces to

$$-\partial_\tau f + K_\mathrm{u} \partial_{r_\mathrm{u}} f + L_\mathrm{u} \left[(\partial_{r_\mathrm{u}} f)^2 + \partial^2_{r_\mathrm{u}} f \right] - r_\mathrm{d} = 0,$$

and if we differentiate the previous equation with respect to r_d, we obtain that $-1 = 0$, which is a contradiction. In the other case, integrating $\partial_{r_\mathrm{d}} f = -D$ with respect to r_d yields the form of the solution f as $f = -D(\tau) r_\mathrm{d} + w(\tau, r_\mathrm{u})$ for some function $w(\tau, r_\mathrm{u})$. By setting such a solution to PDE (2.13), we obtain that $\partial_{r_\mathrm{u}} w = 0$, i.e. $\partial_{r_\mathrm{u}} f = 0$, which leads to a contradiction in the same way as in the previous case. Therefore, the term in (2.15) is not constantly zero, which completes the proof. □

The solution to the two-factor convergence model of CIR type with $\rho = 0$ (and its properties) with Theorem 11 are the main theoretical results of this thesis.

3
Model calibration

In the previous two chapters we introduced a few term structure models of interest rate, and analysed their mathematical properties. However, to put such a model in use one has to estimate its parameters – we call this “the calibration of the model.” The main objective of this chapter is to calibrate selected models: the one-factor Vasicek and CIR models, and the two-factor convergence models of Vasicek and CIR type.

There are a number of works dealing with the estimation of the parameters of the diffusion model

$$\mathrm{d}\boldsymbol{X} = \boldsymbol{\mu}(\boldsymbol{X}, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\boldsymbol{X}, t, \boldsymbol{\theta})\mathrm{d}\boldsymbol{W}, \tag{3.1}$$

based on various techniques, for example, simulation methods (Gouriéroux, Monfort and Renault (1993), Gallant and Tauchen (1996)) or the generalized method of moments (Hansen and Scheinkman (1995), Kessler and Sorensen (1999)). Another approach, which we employ in this thesis, is to find the density (the so-called *transition function*) of a diffusion process and apply a *maximum likelihood estimator*. For instance, consider an equidistant discrete sample $\{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$ of a multivariate $\boldsymbol{\theta}$-parametrized Itō process $\boldsymbol{X}_t$ described by a SDE (3.1), and let $f(\boldsymbol{x}_i, \Delta, \boldsymbol{\theta}|\boldsymbol{x}_{i-1})$ be the density of $\boldsymbol{X}_i$ conditioned on the previous value, where Δ is the time distance between two values. Then it is easy to obtain the corresponding *log-likelihood function*

$$\ell(\boldsymbol{\theta}) = \sum_{i=2}^{n} \ln f(\boldsymbol{x}_i, \Delta, \boldsymbol{\theta}|\boldsymbol{x}_{i-1}),$$

and finding a (global) maximum is the only thing left to do.

In general, one way to find the density of an Itō process is to solve the so-called *Fokker-Planck equation*, and we discuss this possibility in Section 3.2. We show that since solving a Fokker-Planck equation is not a very practical approach, it is only possible to find its analytical solution in special cases. The other (and more practical) way, which we introduce in Section 3.3, is a very recent result written by Aït-Sahalia (2002) and Aït-Sahalia (2008). This approach enables us to find an analytic approximation of the density of $\boldsymbol{X}_t$ and the corresponding log-likelihood function, respectively. However, it is often hard to compute constants in the expansion of the density of $\boldsymbol{X}_t$, especially for multivariate models; nevertheless, we exploit this idea to determine a similar approximation.

We estimate the model using Slovak and EMU's data before Slovak Republic adopted the Euro currency, which we discuss in the following section. Then we give a detailed review of the estimation methods, exhibit the estimates of selected models and compare them.

3.1 The data

We use Slovak and EMU data. In particular, the data consist of 62 daily (that is, $\Delta = 1/252$) observations from 1^{st} October 2008 to 31^{st} December 2008. The reason why we use data for such a short period is the influence of the economic crisis. Figure 3.1 depicts the evolution of the Slovak and EMU's overnight, 1 week, 2 week, 1 month, 2 month, 3 month, 6 month, 9 month and 1 year interest rates from 2^{nd} June 2008 to 31^{st} December 2008. In the last quarter of 2008 (from 1^{st} October 2008) we can see the influence of the upcoming economic crisis that highlights the strong dependence between both interest rates immediately before the Slovak Republic adopted the Euro currency. We note that there is a structural breakpoint, that is, a change in the settings of the economy (and, therefore, in the parameters). Nevertheless, this period offers enough observations to estimate the selected models. The European market data, EONIA and EURIBOR, are available at *http://www.euribor.org*. The Slovak market data BRIBOR are taken as the middle between an offer and a bid, which is available at the National Bank of Slovakia website, *http://www.nbs.sk*. We use the overnight interest rates as the short-rates.

Tables 3.1 and 3.2 provide some descriptive statistics for the short-rates.

3.2 The Fokker-Planck equation and the distribution of an Itō process

Consider that $\{\boldsymbol{X}_t\}$ with values in $\mathbb{R}^n$ is governed by SDE (1.1), and let us denote $\boldsymbol{\Sigma}(\boldsymbol{X}, t, \boldsymbol{\theta}) = \frac{1}{2}\boldsymbol{\sigma}(\boldsymbol{X}, t, \boldsymbol{\theta})\boldsymbol{\sigma}(\boldsymbol{X}, t, \boldsymbol{\theta})'$. Then the density f of process $\boldsymbol{X}_t$ solves the so-called Fokker-Planck equation

$$\partial_t f(\boldsymbol{x}, t, \boldsymbol{\theta}) = -\sum_{i=1}^{n} \partial_{x_i}[\mu_i(\boldsymbol{x}, t, \boldsymbol{\theta}) f(\boldsymbol{x}, t, \boldsymbol{\theta})] + \sum_{i=1}^{n}\sum_{j=1}^{n} \partial^2_{x_i x_j}[\Sigma_{ij}(\boldsymbol{x}, t, \boldsymbol{\theta}) f(\boldsymbol{x}, t, \boldsymbol{\theta})]. \tag{3.2}$$

In the case of a one-dimensional process (3.1) with only one Wiener process increment, the Fokker-Planck equation reduces to

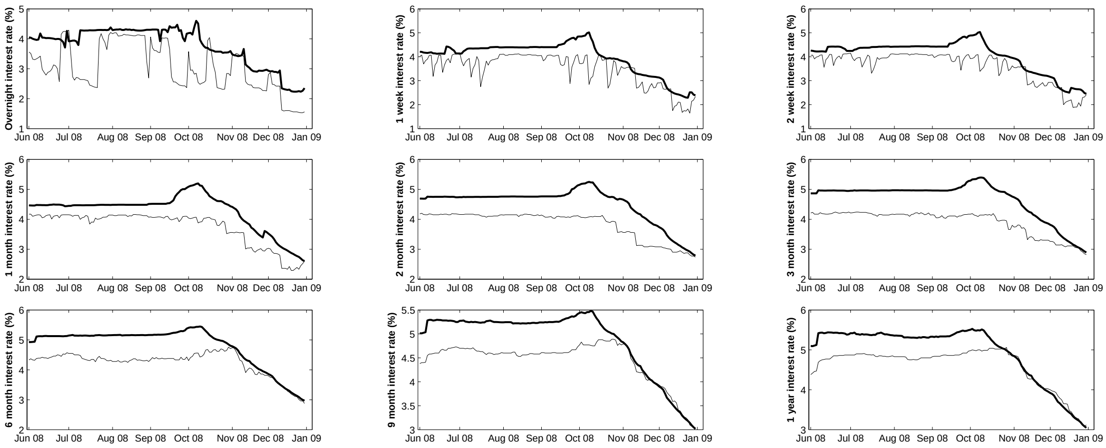

Figure 3.1: *The evolution of BRIBOR (thin line) and EURIBOR (thick line) from* 2^{nd} *June* 2008 *to* 31^{st} *December* 2008*.*

	$r_{\mathbf{SVK}}$	$\mathrm{d}r_{\mathbf{SVK}}$	$r_{\mathbf{EMU}}$	$\mathrm{d}r_{\mathbf{EMU}}$	$r_{\mathbf{EMU}} - r_{\mathbf{SVK}}$
Mean	0.025258	-0.000330	0.031916	-0.000301	0.006658
Std. dev.	0.006589	0.003047	0.006322	0.001359	0.004772
Skewness	0.196581	1.069584	0.098701	-2.523980	0.283366
Kurtosis	2.478430	9.940486	2.316502	11.911318	3.252789
Maximum	0.040400	0.013400	0.046000	0.002500	0.019700
Minimum	0.015300	-0.008400	0.022340	-0.005940	-0.004000

Table 3.1: *Individual descriptive statistics for the selected variables.*

	$\mathrm{d}r_{\mathbf{SVK}}$	$r_{\mathbf{EMU}}$	$\mathrm{d}r_{\mathbf{EMU}}$	$r_{\mathbf{EMU}} - r_{\mathbf{SVK}}$
$r_{\mathbf{SVK}}$	0.227530	0.715676	0.101294	-0.423370
$\mathrm{d}r_{\mathbf{SVK}}$		0.001662	0.318626	-0.305394
$r_{\mathbf{EMU}}$			0.079707	0.329754
$\mathrm{d}r_{\mathbf{EMU}}$				-0.033528

Table 3.2: *Correlation of the selected variables.*

$$\partial_t f(x,t,\boldsymbol{\theta}) = -\partial_x[\mu(x,t,\boldsymbol{\theta}) f(x,t,\boldsymbol{\theta})] + \frac{1}{2}\partial_x^2[\sigma^2(x,t,\boldsymbol{\theta}) f(x,t,\boldsymbol{\theta})]. \tag{3.3}$$

The stationary density can be obtained by setting $\partial_t f(\boldsymbol{x}, t, \boldsymbol{\theta}) = 0$. Note that the Fokker-Planck equation can only be solved analytically in special cases. In the following we offer a solution to the one-factor Vasicek model mentioned in the previous chapters, which demonstrates that solving the Fokker-Planck equation for a very simple model is rather complicated.

Example: the density of the one-factor Vasicek model

In the case of the one-factor Vasicek model (1.3) the corresponding Fokker-Planck equation is

$$\partial_t f = -\partial_r[\kappa(\theta - r)f] + \frac{\sigma^2}{2}\partial_r^2 f = \kappa f - \kappa\theta\partial_r f + \kappa r \partial_r f + \frac{\sigma^2}{2}\partial_r^2 f.$$

The idea is to find the *characteristic function* of r_t (denoted by $\psi = \psi(s,t)$), which defines the distribution of r_t uniquely. Therefore, we multiply the equation by $\exp\{\mathrm{i}sr\}$, and integrate on $\mathbb{R}$ with respect to r. We obtain that

$$\begin{aligned}\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_t f)\mathrm{d}r &= \kappa\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}f)\mathrm{d}r - \kappa\theta\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_r f)\mathrm{d}r \\ &+ \kappa\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}r\partial_r f)\mathrm{d}r + \frac{\sigma^2}{2}\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_r^2 f)\mathrm{d}r.\end{aligned}$$

Since $\lim_{r\to\pm\infty} f = \lim_{r\to\pm\infty}\partial_r f = 0$, it is easy to see that

$$\begin{aligned}\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_t f)\mathrm{d}r &= \partial_t\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}f)\mathrm{d}r = \partial_t\psi, \\ \int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_r f)\mathrm{d}r &= \exp\{\mathrm{i}sr\}f\Big|_{-\infty}^{\infty} - \mathrm{i}s\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}f)\mathrm{d}r = -\mathrm{i}s\psi, \\ \int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_r^2 f)\mathrm{d}r &= \exp\{\mathrm{i}sr\}\partial_r f\Big|_{-\infty}^{\infty} - \mathrm{i}s\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_r f)\mathrm{d}r = -s^2\psi, \\ \int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}r\partial_r f)\mathrm{d}r &= \frac{1}{\mathrm{i}}\partial_s\int_{\mathbb{R}}(\exp\{\mathrm{i}sr\}\partial_r f)\mathrm{d}r = -\partial_s(s\psi) = -(\psi + s\partial_s\psi),\end{aligned}$$

where we exploited that $\exp\{\mathrm{i}sr\}r = \partial_s\exp\{\mathrm{i}sr\}/i$. Consequently

$$\partial_t\psi = \kappa\psi + \kappa\theta\mathrm{i}s\psi - \kappa(\psi + s\partial_s\psi) - \frac{\sigma^2}{2}s^2\psi,$$

so, the characteristic function of r_t satisfies the quasi-linear PDE

$$\partial_t\psi + \kappa s\partial_s\psi = (\kappa\theta\mathrm{i}s - \frac{\sigma^2}{2}s^2)\psi. \tag{3.4}$$

We can solve this equation using the method of characteristics (see, i.e., Evans (1998) or Ševčovič (2008)). Let $\phi = \phi(t,s,\psi)$ be a corresponding auxiliary function, that is, ϕ is a solution to

$$\partial_t\phi + \kappa s\partial_s\phi + \psi\left(\mathrm{i}\kappa\theta - \frac{\sigma^2}{2}s^2\right)\partial_\psi\phi = 0.$$

Then, the solution ψ of the original problem (3.4) satisfies $\phi[t,s,\psi(s,t)] = 0$. The corresponding characteristics are

$$\begin{aligned}\dot{t} &= 1, \\ \dot{s} &= \kappa s, \\ \dot{\psi} &= \psi\left(\mathrm{i}\kappa\theta - \frac{\sigma^2}{2}s^2\right).\end{aligned}$$

The second equation implies that $\mathrm{d}_\tau \ln(s) = \kappa = \kappa\dot{t}$. Consequently, $\mathrm{d}_\tau[\ln(s) - \kappa t] = \mathrm{d}_\tau[\ln(s) - \ln(\exp\{\kappa t\})] = \mathrm{d}_\tau[\ln(s\exp\{-\kappa t\})] = 0 = \mathrm{d}_\tau(s\exp\{-\kappa t\})$. The second equation also implies that $\frac{\sigma^2}{2}s^2 = \frac{\sigma^2}{2\kappa}s\dot{s} = \frac{\sigma^2}{4\kappa}\mathrm{d}_\tau(s^2)$. Therefore, $\mathrm{d}_\tau \ln(\psi) = \dot{\psi}/\psi = \mathrm{i}\theta\dot{s} - \frac{\sigma^2}{4\kappa}\mathrm{d}_\tau(s^2)$. Hence, a solution $\psi(s,t)$ to the original problem (3.4) is any solution to the equation

$$\phi(t,s,\psi) = \Phi\left(s\exp\{-\kappa t\}, \ln(\psi) + \frac{\sigma^2}{4\kappa}s^2 - \mathrm{i}\theta s\right) = 0, \tag{3.5}$$

where $\Phi(\xi_1,\xi_2)$ is continuously differentiable. The initial condition ($t = 0$), which has to be satisfied,

$$\psi(s,0) = \int_{\mathbb{R}} \exp\{\mathrm{i}sr_0\}\delta(r - r_0)\mathrm{d}r = \exp\{\mathrm{i}sr_0\},$$

implies

$$\phi(0,s,\psi(s,0)) = \Phi\left(s, \mathrm{i}sr_0 + \frac{\sigma^2}{4\kappa}s^2 - \mathrm{i}\theta s\right) = 0.$$

Consequently, $\Phi(\xi_1,\xi_2) = \mathrm{i}r_0\xi_1 + \frac{\sigma^2}{4\kappa}\xi_1^2 - \mathrm{i}\theta\xi_1 - \xi_2$. By applying this to the solution (3.5) we obtain that

$$\psi(s,t) = \exp\left\{\mathrm{i}s[\theta(1 - \exp\{-\kappa t\}) + r_0\exp\{-\kappa t\}] - \frac{1}{2}s^2\left[\frac{\sigma^2}{2\kappa}(1 - \exp\{-2\kappa t\})\right]\right\},$$

which is nothing else but the characteristic function of the normal distribution with the mean value $\theta(1 - \exp\{-\kappa t\}) + r_0\exp\{-\kappa t\}$ and variance $\frac{\sigma^2}{2\kappa}(1 - \exp\{-2\kappa t\})$. Therefore,

$$r_t|r_0 \sim \mathcal{N}\left(\theta(1 - \exp\{-\kappa t\}) + r_0\exp\{-\kappa t\}, \frac{\sigma^2}{2\kappa}(1 - \exp\{-2\kappa t\})\right).$$

For $t \to \infty$ we obtain (the same result if we solved the stationary Fokker-Planck equation) that $r_t \sim \mathcal{N}(\theta, \frac{\sigma^2}{2\kappa})$.

Example: the density of the one-factor CIR model

Another important example is the density function of the one-factor CIR model, although, we do not offer it as a solution to the corresponding Fokker-Planck equation. The density (Cox et al. (1985)) is

$$f(r_t|r_0) = c\exp\{-u - v\}(v/u)^{q/2}\mathrm{I}_q\left(2\sqrt{uv}\right), \tag{3.6}$$

where

$$\begin{aligned}
c &= \frac{2\kappa}{\sigma^2(1 - \exp\{-\kappa t\})},\\
u &= cr_0\exp\{-\kappa t\},\\
v &= cr_t,\\
q &= \frac{2\kappa\theta}{\sigma^2} - 1,
\end{aligned}$$

and I_q is the modified Bessel function of the first kind of order q. Equation (3.6) can be rewritten in the form

$$f(r_t|r_0) = 2cf_{\chi'^2}(2v, 2u, 2q+2),$$

where $f_{\chi'^2}(x, \lambda, k)$ is the density function of the non-central χ^2 distribution with k degrees of freedom and non-centrality parameter λ. That is, $2cr_t$ has the non-central χ^2 distribution with $2q+2$ degrees of freedom and non-centrality parameter $2u$.

3.3 An analytic approximation of the density of a univariate Itō process

As we have mentioned at the beginning of this chapter, Aït-Sahalia (2002) offers an alternative to solving a Fokker-Planck equation: a closed form approximation of the density of a diffusion process. In the sequel we introduce this approach for a process that follows SDE (3.1).

According to Aït-Sahalia (2002), the idea is analogous to the Central Limit Theorem. Let us consider a sample size n of standardized random variables. Then the sum of these variables can be approximated by a limiting normal distribution. If the sample is not large enough, higher order terms (in an expansion, which we discuss later) can be calculated to improve the small sample performance. In this analogy the process $\{X_t\}$ represents a sum of random variables, and Δ represents the size of the sample n ($\Delta \to 0$ corresponds to $n \to \infty$).

First, we standardize X_t using the transformation

$$Y_t = \gamma(X_t, \boldsymbol{\theta}), \text{ where } \gamma(x, \boldsymbol{\theta}) = \int \frac{\mathrm{d}x}{\sigma(x, \boldsymbol{\theta})},$$

i.e. γ is any primitive function of $1/\sigma$ (the constant is irrelevant). By applying Itō's lemma we obtain that

$$\mathrm{d}Y = \mu_Y(Y, \boldsymbol{\theta})\mathrm{d}t + \mathrm{d}W, \text{ where } \mu_Y(y, \boldsymbol{\theta}) = \frac{\mu(\gamma^{\mathrm{inv}}(y, \boldsymbol{\theta}), \boldsymbol{\theta})}{\sigma(\gamma^{\mathrm{inv}}(y, \boldsymbol{\theta}), \boldsymbol{\theta})} - \frac{1}{2}\partial_x\sigma(\gamma^{\mathrm{inv}}(y, \boldsymbol{\theta}), \boldsymbol{\theta}),$$

that is, Y_t has a unit diffusion. Therefore, we can understand this as Y being "closer" to a normal distribution than X because of its standardization. The relation between f and f_Y is

$$f(x, \Delta, \boldsymbol{\theta}|X_0) = [\sigma(x, \boldsymbol{\theta})]^{-1} f_Y[\gamma(x, \boldsymbol{\theta}), \Delta, \boldsymbol{\theta}|\gamma(X_0, \boldsymbol{\theta})]. \tag{3.7}$$

However, the density f_Y of Y_t becomes peaked around the conditional value Y_0 when Δ becomes small, so it is not suitable for expansion. Therefore, we make another transformation

$$Z_t = \Delta^{-1/2}(Y_t - Y_0),$$

which is an analogy to centring the sum of the standardized variables and dividing it by $n^{1/2}$. For a fixed Δ, the distribution of Z is close enough to the standard normal distribution, so it is possible to make a convergent expansion around the $\mathcal{N}(0,1)$ term. The density of Z is $f_Z(z,\Delta,\boldsymbol{\theta}|Y_0) = \Delta^{1/2} f_Y(\Delta^{1/2} z + Y_0, \Delta, \boldsymbol{\theta}|Y_0)$. After we have obtained a sequence of the approximations to f_Z, we make an inverse transformation, and we have

$$f(x,\Delta,\boldsymbol{\theta}|X_0) = [\sigma(x,\boldsymbol{\theta})]^{-1} \times \Delta^{-1/2} f_Z(\Delta^{-1/2}[\gamma(x,\boldsymbol{\theta}) - \gamma(X_0,\boldsymbol{\theta})], \Delta, \boldsymbol{\theta}|\gamma(X_0,\boldsymbol{\theta})).$$

To approximate the density of Z we use a Hermite series expansion. The Hermite polynomials are

$$H_j(z) = [\phi(z)]^{-1} \frac{\mathrm{d}^j \phi(z)}{\mathrm{d}z^j},\ j \geq 0,\ \text{where } \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\{-z^2/2\},$$

and the expansion is

$$f_Z^{(J)}(z,\Delta,\boldsymbol{\theta}|Y_0) = \phi(z) \sum_{j=0}^{J} \eta_Z^{(j)}(\Delta,\boldsymbol{\theta}|Y_0) H_j(z), \tag{3.8}$$

where

$$\begin{aligned} \eta_Z^{(j)}(\Delta,\boldsymbol{\theta}|Y_0) &= \frac{1}{j!} \int_{\mathbb{R}} H_j(z) f_Z(z,\Delta,\boldsymbol{\theta}|Y_0) \mathrm{d}z \\ &= \frac{1}{j!} \mathrm{E}[H_j(\Delta^{-1/2}(Y_{t+1} - Y_0)), \boldsymbol{\theta}|Y_t = Y_0]. \end{aligned}$$

It is possible to approximate these coefficients in a closed form by applying Taylor's Theorem to the function

$$s \mapsto \mathrm{E}[g(Y_{t+s}, Y_0)|Y_t = Y_0],$$

which implies

$$\mathrm{E}[g(Y_{t+\Delta}, Y_0)|Y_t = Y_0] \approx \sum_{k=0}^{K} A^k(\boldsymbol{\theta}) \cdot g(Y_0, Y_0) \frac{\Delta^k}{k!},$$

where $A(\boldsymbol{\theta})$ is an operator defined by

$$A(\boldsymbol{\theta}) : g \mapsto \mu_Y(.,\boldsymbol{\theta}) \partial_y g + \frac{1}{2} \partial_y^2 g.$$

Once we are given an approximation $f_Z^{(J)}$ (or better $f_Z^{(J,K)}$), we apply the inverse transformations to obtain the approximation $f_X^{(J)}$ to f_X.

There is also another way to perform this approximation. Instead of increasing the order of the Hermite polynomials, we can increase the powers of Δ, that is, we let $f_Z^{(K)} = f_Z^{(\infty,K)}$. We obtain an explicit representation of $f_Y^{(K)}$

$$f_Y^{(K)}(y,\Delta,\boldsymbol{\theta}|Y_0) = \Delta^{-1/2} \phi\left(\frac{y - Y_0}{\Delta^{1/2}}\right) \exp\left\{\int_{Y_0}^{y} \mu_Y(w,\boldsymbol{\theta}) \mathrm{d}w\right\} \sum_{k=0}^{K} c_k(y,\boldsymbol{\theta}|Y_0) \frac{\Delta^k}{k!}, \tag{3.9}$$

where

$$\begin{aligned} c_0(y, \boldsymbol{\theta}|Y_0) &= 1, \\ c_j(y, \boldsymbol{\theta}|Y_0) &= j(y - Y_0)^{-j} \\ &\times \int_{Y_0}^{y} (w - Y_0)^{j-1} \left[\lambda_Y(w, \boldsymbol{\theta}) c_{j-1}(w, \boldsymbol{\theta}|Y_0) + \frac{1}{2} \partial_w^2 c_{j-1}(w, \boldsymbol{\theta}|Y_0) \right] \mathrm{d}w, \end{aligned}$$

and

$$\lambda_Y(y, \boldsymbol{\theta}) = -\frac{1}{2}[\mu_Y(y, \boldsymbol{\theta})^2 + \partial_y \mu_Y(y, \boldsymbol{\theta})].$$

By using relation (3.7) we obtain an approximation of the density of the original process.

The assumptions, the conditions of convergence and the asymptotic properties are not discussed; the reader is referred to Aït-Sahalia (2002) for more detail. In the following we offer two examples of analytical approximations, particularly, the one-factor Vasicek and CIR models, and we estimate the corresponding parameters for BRIBOR and EONIA. The next two examples of an analytical approximation of an interest rate model, which we apply to real data, are taken from Aït-Sahalia (1999).

Example: an approximation of the density of the one-factor Vasicek model

Consider the well-known Vasicek model $\mathrm{d}r = \kappa(\theta - r)\mathrm{d}t + \sigma \mathrm{d}W$, where κ, θ and σ are the parameters of the model (i.e., $\boldsymbol{\theta} = (\kappa, \theta, \sigma)'$). An approximation can be obtained by setting

$$\begin{aligned} \gamma(r, \boldsymbol{\theta}) &= r/\sigma, \\ f_Y(y, \Delta, \boldsymbol{\theta}|Y_0) &= f_Y^0(y, \Delta, \boldsymbol{\theta}|Y_0)[1 + c_1(y, \boldsymbol{\theta}|Y_0)\Delta + c_2(y, \boldsymbol{\theta}|Y_0)\Delta^2/2], \\ \text{where} & \\ f_Y^0(y, \Delta, \boldsymbol{\theta}|Y_0) &= \frac{1}{\sqrt{2\pi\Delta}} \exp\left\{ -\frac{(y - Y_0)^2}{2\Delta} - \frac{\kappa(y^2 - Y_0^2)}{2} + \frac{\kappa\theta(y - Y_0)}{\sigma} \right\}, \\ c_1(y, \boldsymbol{\theta}|Y_0) &= -\frac{\kappa}{6\sigma^2}[3\kappa\theta^2 - 3(y + Y_0)\kappa\theta\sigma + \sigma^2(-3 + y^2\kappa + yY_0\kappa + Y_0^2\kappa)], \\ c_2(y, \boldsymbol{\theta}|Y_0) &= \frac{\kappa^2}{36\sigma^4}[9\kappa^2\theta^4 - 18y\kappa^2\theta^3\sigma + 3\kappa\theta^2\sigma^2(-6 + 5y^2\kappa) \\ &\quad -6y\kappa\theta\sigma^3(-3 + y^2\kappa) + \sigma^4(3 - 6y^2\kappa + y^4\kappa^2) \\ &\quad +2Y_0\kappa\sigma(-3\theta + y\sigma)(3\kappa\theta^2 - 3y\kappa\theta\sigma + \sigma^2(-3 + y^2\kappa)) \\ &\quad +3Y_0^2\kappa\sigma^2(5\kappa\theta^2 - 4y\kappa\theta\sigma + \sigma^2(-2 + y^2\kappa)) \\ &\quad +2Y_0^3\kappa^2\sigma^3(-3\theta + y\sigma) + Y_0^4\kappa^2\sigma^4], \end{aligned}$$

to equation (3.7).

Example: an approximation of the density of the one-factor CIR model

The second example is the CIR model $\mathrm{d}r = \kappa(\theta - r)\mathrm{d}t + \sigma\sqrt{r}\mathrm{d}W$, where κ, θ and σ are the parameters of the model. In this case an analytic approximation can be obtained by setting

$$\begin{aligned}
\gamma(r, \boldsymbol{\theta}) &= 2\sigma^{-1}\sqrt{r}, \\
f_Y(y, \Delta, \boldsymbol{\theta}|Y_0) &= f_Y^0(y, \Delta, \boldsymbol{\theta}|Y_0)[1 + c_1(y, \boldsymbol{\theta}|Y_0)\Delta + c_2(y, \boldsymbol{\theta}|Y_0)\Delta^2/2], \\
\text{where} & \\
f_Y^0(y, \Delta, \boldsymbol{\theta}|Y_0) &= \frac{1}{\sqrt{2\pi\Delta}} \exp\left\{-\frac{(y - Y_0)^2}{2\Delta} - \frac{\kappa(y - Y_0)}{4}\right\}(y/Y_0)^{-1/2 + 2\kappa\theta/\sigma^2}, \\
c_1(y, \boldsymbol{\theta}|Y_0) &= -\frac{1}{24yY_0\sigma^4}[48\kappa^2\theta^2 - 48\kappa\theta\sigma^2 + 9\sigma^4 + yY_0\kappa^2\sigma^2(-24\theta + y^2\sigma^2) \\
&\quad + y^2Y_0^2\kappa^2\sigma^4 + yY_0^3\kappa^2\sigma^4], \\
c_2(y, \boldsymbol{\theta}|Y_0) &= \frac{1}{576y^2Y_0^2\sigma^8}[9(256\kappa^4\theta^4 - 512\kappa^3\theta^3\sigma^2 + 224\kappa^2\theta^2\sigma^4 + 32\kappa\theta\sigma^6 \\
&\quad - 15\sigma^8) + 6yY_0\kappa^2\sigma^2(-24\theta + y^2\sigma^2)(16\theta^2\kappa^2 - 16\kappa\theta\sigma^2 + 3\sigma^4) \\
&\quad + y^2Y_0^2\kappa^2\sigma^4(672\kappa^2\theta^2 - 48\kappa\theta\sigma^2(2 + y^2\kappa) + \sigma^4(-6 + y^4\kappa^2)) \\
&\quad + 2yY_0^3\kappa^2\sigma^4(48\kappa^2\theta^2 - 24\kappa\theta\sigma^2(2 + y^2\kappa) + \sigma^4(9 + y^4\kappa^2)) \\
&\quad + 3y^2Y_0^4\kappa^4\sigma^6(-16\theta + y^2\sigma^2) + 2y^3Y_0^5\kappa^4\sigma^8 + y^2Y_0^6\kappa^4\sigma^8].
\end{aligned}$$

to equation (3.7).

3.4 Estimates in one-factor models and pitfalls of the approximation approach

In this section we employ a maximum likelihood estimator to estimate the parameters of the one-factor Vasicek and CIR models. For both the BRIBOR and EONIA datasets, we find the maximum likelihood estimates using the true and approximate densities. All the computations were done in the *Wolfram Mathematica* environment.

There are two important aspects in the maximization of a corresponding approximate (log-)likelihood function: the settings of the selected algorithm and the "stability" of an approximate density function. For instance, an application of the global maximization procedure to the log-likelihood (usually the simulated annealing or any evolution algorithm) might not behave well. Surprisingly, when we applied this strategy to the log-likelihood functions using the true densities (and approximate densities as well), the procedure failed. On the other hand, application of the local maximization procedure with a suitable initial point returned "wise" results. The "stability" of the approximate density is a very important aspect. A good example is the approximation of the density of the one-factor CIR model, which we offer in Section 3.3. By setting $\kappa = 20$ and $\sigma = 0.05$, we can see that there is no approximation, and it does not matter if we employ approximation (3.8) or (3.9) (see Figure

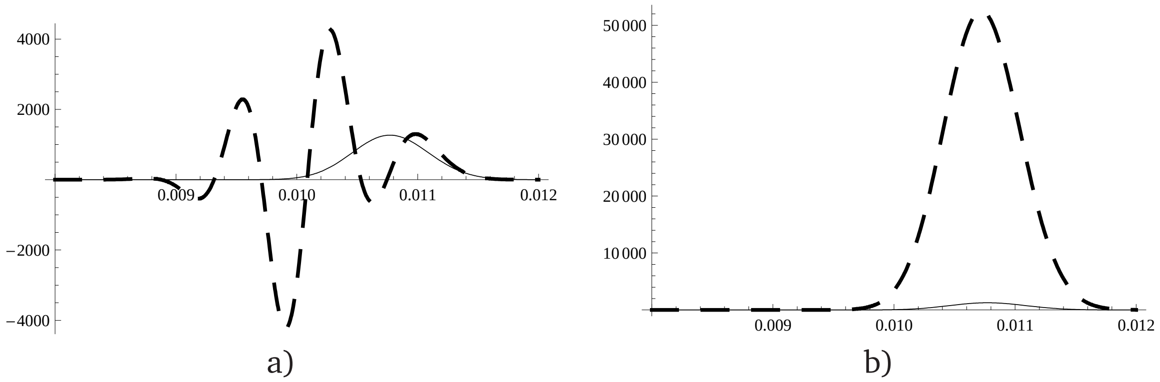

Figure 3.2: *A problem with an approximate density in the one-factor CIR model with parameters* $\kappa = 10$, $\theta = 0.03$, $\sigma = 0.01$ *and starting point* $r_0 = 0.01$ *in the case of using the approximation a)* (3.8) *and b)* (3.9).

	BRIBOR			**EONIA**		
	κ	θ	σ	κ	θ	σ
T	29.7544	0.0224798	0.0491643	7.53173	0.0219006	0.0214246
A	29.7431	0.0224763	0.0491624	7.55337	0.0219156	0.0214240
INIT	20.0000	0.0200000	0.0500000	2.00000	0.0200000	0.0500000

Table 3.3: *Maximum likelihood estimates of the parameters in the one-factor Vasicek model for the BRIBOR and EONIA data using the true (T) and approximate (A) densities.*

	BRIBOR			**EONIA**		
	κ	θ	σ	κ	θ	σ
T	25.494	0.0220148	0.288818	7.74489	0.022173	0.117199
A	25.486	0.0220165	0.288821	7.75816	0.022182	0.117200
INIT	20.000	0.0200000	0.500000	7.70000	0.022000	0.117000

Table 3.4: *Maximum likelihood estimates of the parameters in the one-factor CIR model for the BRIBOR and EONIA data using the true (T) and approximate (A) densities.*

3.2). Moreover, the difference was increasing in κ. Therefore, it is reasonable to use the local maximization algorithm with a properly selected initial point; nevertheless, no one can ensure that the true estimate is not equal to an improper point.

Tables 3.3 and 3.4 exhibit maximum likelihood estimates in the one-factor Vasicek and CIR models, respectively. It is clear that there is almost no difference in the estimates. Consequently, we accept the estimates obtained by using the analytic approximation approach. We employed the algorithm for finding the local maxima with an initial point INIT. The initial point in the case of estimation parameters in the one-factor Vasicek model did not seem to influence the estimation. We ran the optimization for different initial points, and we obtained the same estimates for each run. According to the foregoing result of the estimation, we offer $\kappa_0 = 20$, $\theta_0 = 0.02$ and $\sigma_0 = 0.05$ to be a good initial point for BRIBOR and $\kappa_0 = 2$, $\theta_0 = 0.02$ and $\sigma_0 = 0.05$ to be a good initial point for EONIA. In the case of the estimation

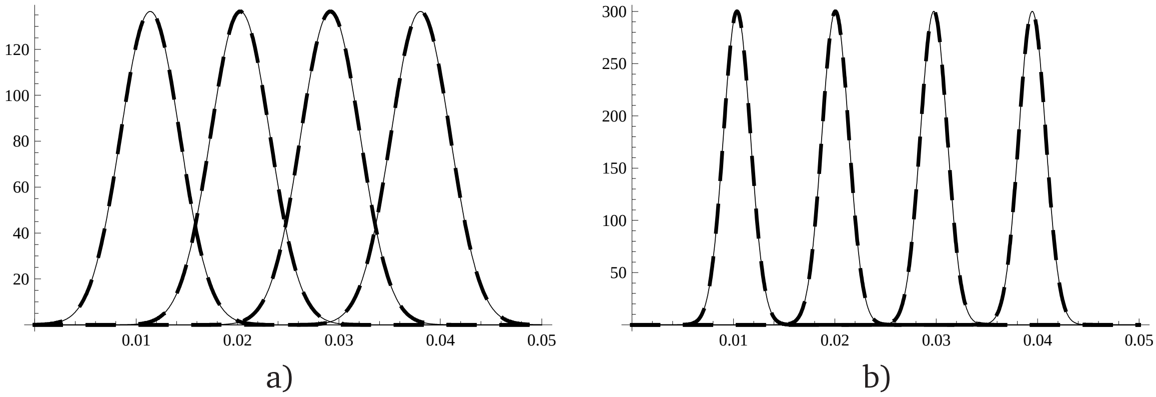

Figure 3.3: *Estimated densities of the one-factor Vasicek model for a) BRIBOR b) EONIA, with the initial value r_0 equal to $0.01, 0.02, 0.03, 0.04$. The solid line corresponds to the true density, and the bold dashed line corresponds to the approximate density.*

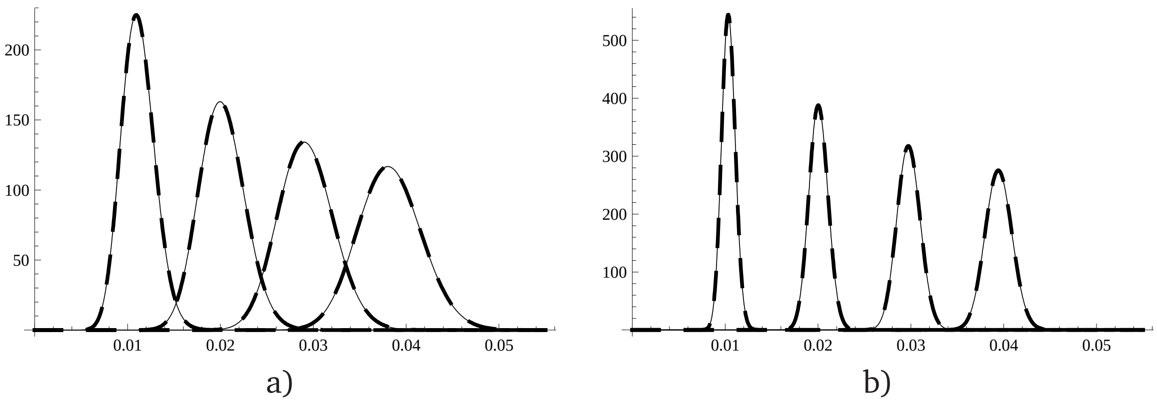

Figure 3.4: *Estimated densities of the one-factor CIR model for a) BRIBOR b) EONIA, with initial value r_0 equal to $0.01, 0.02, 0.03, 0.04$. The solid line corresponds to the true density and the bold dashed line corresponds to the approximate density.*

parameters in the one-factor CIR model, the situation was different. Although the estimation ran well for the BRIBOR data (we exploited the initial point offered in the Vasicek model with a rescaled σ), we faced the problem of the convergence of the local maximization algorithm for the EONIA data; namely, there was a problem with convergence when employing the true density. So the initial point was based on the estimate using the approximate density in the CIR model for the EMU data. In Figures 3.3 and 3.4 we offer a visual comparison of the estimated Vasicek and CIR models' densities for the Slovak and EMU's data.

In the following we estimate the market prices of risk in both models using the real market data. Since the term structures of interest rates developed from the both one-factor Vasicek and CIR models are known (see Section 1.3), we estimate the market prices of risk by minimizing the *weighted mean square error* (WMSE). Assume that each day $i = 1, \dots, I$ (in chronological order) we are given the values of the interest rates for periods of the length τ_j, $j = 1 \dots, J$ (the term structure given by market data). Therefore, every value of the interest rate can be uniquely

	BRIBOR		EURIBOR	
	$\boldsymbol{\lambda}, \boldsymbol{\nu}$	WMSE	$\boldsymbol{\lambda}, \boldsymbol{\nu}$	WMSE
Vasicek	-10.8860	0.0141260	-9.1892	0.0152413
CIR	-43.5283	0.0112398	-38.919	0.0140444

Table 3.5: *Market prices of risk estimates.*

identified by day and period, and we denote its value by R_{ij}. Then

$$\mathrm{WMSE}(\lambda) = \sum_{j=1}^{J} \sum_{i=1}^{I} W_{ij} [R(r_i, \tau_j, \lambda) - R_{ij}]^2, \tag{3.10}$$

where W_{ij}, $i = 1, \ldots, I$, $j = 1, \ldots, J$, are weights, and $R(r_i, \tau_j \lambda)$ is the predicted term structure depending on the overnight interest rate r_i, period τ_j, and the market price of the risk. We note that the market price of risk in the Vasicek model is equal to a constant λ, but in the case of the CIR model we have $\lambda(r) = \nu r^{1/2}$; therefore, we minimize the WMSE through ν. The choice of the weights W_{ij} is determined by our preferences; nevertheless, for fixed j and $i_1 < i_2$ the condition $W_{i_1 j} \leq W_{i_2 j}$ has to be satisfied (it is reasonable to put more weight on more recent data). For our purpose we used uniform weighing (i.e., we set $W_{ij} = 1$ for all i and j), so we use the same error evaluation as Corzo and Schwartz (2000) did. Other choice of the weights could be, for instance, $W_{ij} = \tau_j^2$ (Ševčovič and Urbánová-Csajková (2005)). Another approach for estimating the market price of risk is to minimize the WMSE for the predicted and true bond price. However, we are primarily given the market term structures of interest rates; therefore, we minimized the WMSE for the terms structures of interest rates.

In Figures 3.5 and 3.6 we illustrate typical market term structures compared with the theoretical term structures for the Vasicek and CIR models.

3.5 An analytic approximation of the density of a multivariate Itō process

Aït-Sahalia (2008) employed the idea presented in his paper Aït-Sahalia (2002) to develop an approximate log-likelihood function for multivariate processes. However, higher order terms in this approximation are (depending on the model) difficult to compute. We offer a similar method, based on the previously mentioned papers, to figure out an approximate density of a multivariate process using a multivariate normal density directly, without computing any expansion terms.

Let $\boldsymbol{X}_t$, with values in $\mathbb{R}^n$ and an initial value $\boldsymbol{X}_0$, follow SDE (3.1) and let $\boldsymbol{\gamma}(\boldsymbol{x}, \boldsymbol{\theta})$ be an invertible function such that

$$\mathbf{J}_{\boldsymbol{x}} \boldsymbol{\gamma}(\boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{\sigma}^{-1}(\boldsymbol{x}, \boldsymbol{\theta}), \tag{3.11}$$

where $\boldsymbol{\sigma}^{-1}(\boldsymbol{x}, \boldsymbol{\theta})$ is the matrix inverse of $\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{\theta})$. By $\boldsymbol{\gamma}^{\mathrm{inv}}(., \boldsymbol{\theta})$ we denote the inverse transformation for $\boldsymbol{\gamma}(., \boldsymbol{\theta})$. In this case we say that the diffusion $\boldsymbol{X}_t$ is reducible

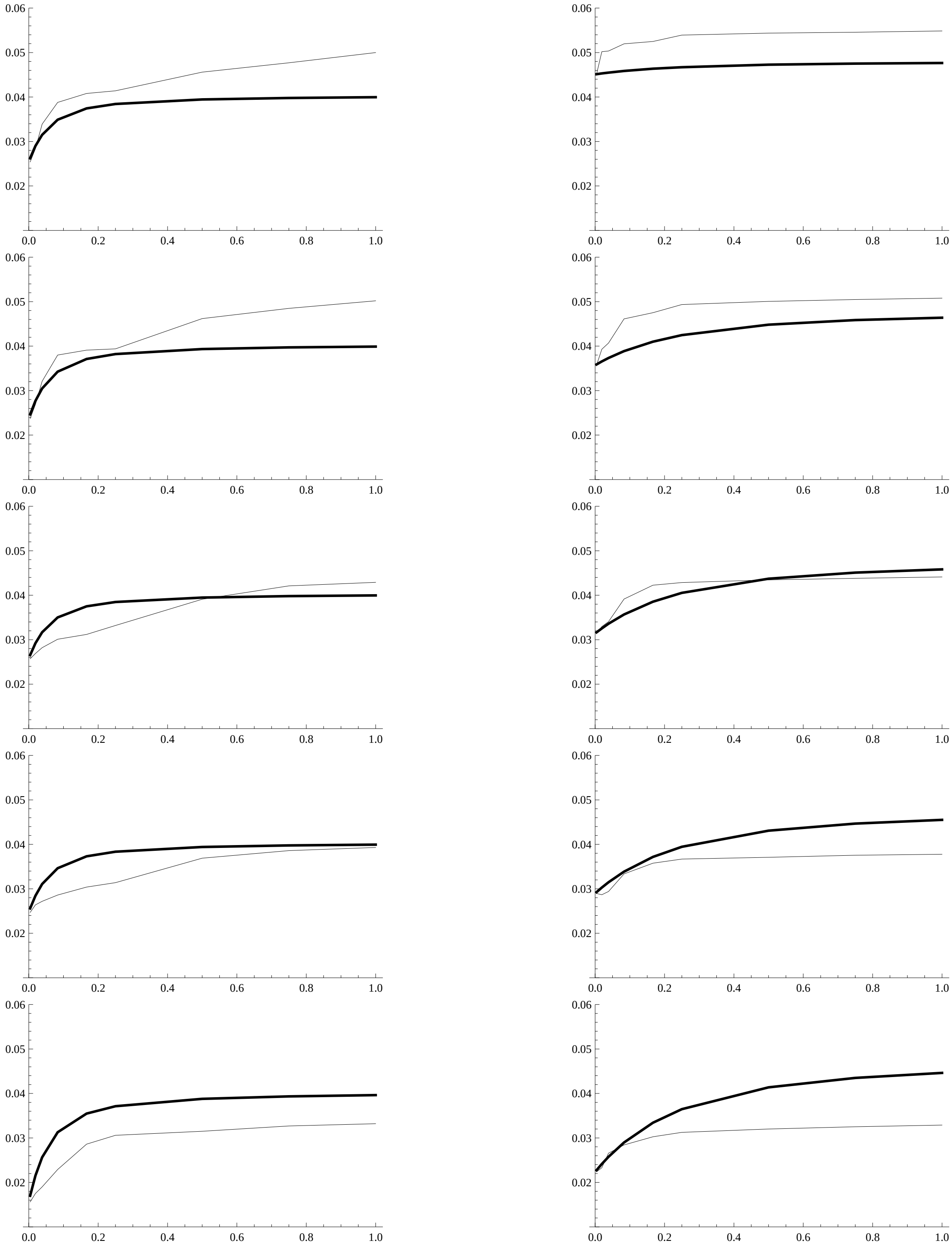

Figure 3.5: *Theoretical term structures of interest rate (thick line) in the one-factor Vasicek model for BRIBOR (left column) and EURIBOR (right column) compared with the market term structures (thin line).*

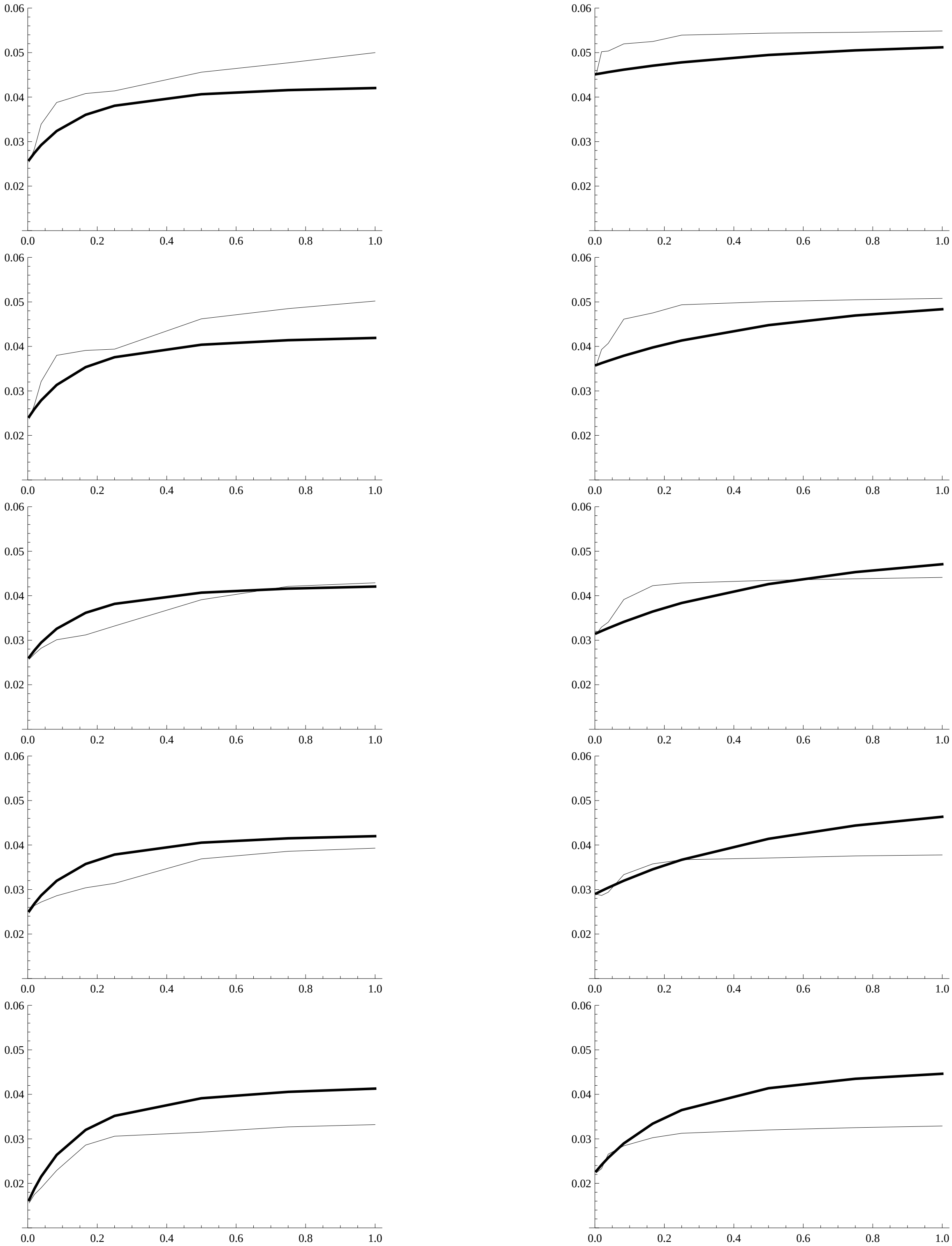

Figure 3.6: *Theoretical term structures of interest rate (thick line) in the one-factor CIR model for BRIBOR (left column) and EURIBOR (right column) compared with the market term structures (thin line).*

to a unit diffusion (as we will see later, our models are reducible diffusions). Itō's lemma implies that the process $\boldsymbol{Y}_t = \boldsymbol{\gamma}(\boldsymbol{X}_t, \boldsymbol{\theta})$ follows SDE

$$\mathrm{d}\boldsymbol{Y} = \boldsymbol{\mu}_{\mathbf{Y}}(\boldsymbol{Y})\mathrm{d}t + \mathrm{d}\boldsymbol{W},$$

where

$$\boldsymbol{\mu}_{\mathbf{Y}}(\boldsymbol{y})\mathrm{d}t = \boldsymbol{\sigma}^{-1}[\boldsymbol{\gamma}^{\mathrm{inv}}(\boldsymbol{y}), \boldsymbol{\theta}]\boldsymbol{\mu}[\boldsymbol{\gamma}^{\mathrm{inv}}(\boldsymbol{y}), \boldsymbol{\theta}]\mathrm{d}t + \left.\begin{pmatrix} (\mathrm{d}\boldsymbol{X})'(\boldsymbol{\nabla}^2_{\boldsymbol{x}}\gamma_1(\boldsymbol{x}, \boldsymbol{\theta}))\mathrm{d}\boldsymbol{X} \\ \vdots \\ (\mathrm{d}\boldsymbol{X})'(\boldsymbol{\nabla}^2_{\boldsymbol{x}}\gamma_n(\boldsymbol{x}, \boldsymbol{\theta}))\mathrm{d}\boldsymbol{X} \end{pmatrix}\right|_{\boldsymbol{x}=\boldsymbol{\gamma}^{\mathrm{inv}}(\boldsymbol{y},\boldsymbol{\theta})}, \tag{3.12}$$

$\gamma_i(\boldsymbol{x}, \boldsymbol{\theta})$ is the ith component of $\boldsymbol{\gamma}(\boldsymbol{x}, \boldsymbol{\theta})$, $(\mathrm{d}W_i)^2 = \mathrm{d}t$ and $(\mathrm{d}t)^2 = \mathrm{d}t\mathrm{d}W_i = 0$. Clearly, $\boldsymbol{Y}_0 = \boldsymbol{\gamma}(\boldsymbol{X}_0, \boldsymbol{\theta})$. Now, we perform another transformation, namely $\boldsymbol{Z}_t = \Delta^{-1/2}(\boldsymbol{Y}_t - \boldsymbol{Y}_0)$, and we recall that Δ is the time difference between two values in the given sample. Again, Itō's lemma yields

$$\mathrm{d}\boldsymbol{Z} = \Delta^{-1/2}\mathrm{d}\boldsymbol{Y} = \Delta^{-1/2}\boldsymbol{\mu}_{\mathbf{Y}}(\boldsymbol{Y}, \boldsymbol{\theta})\mathrm{d}t + \Delta^{-1/2}\mathrm{d}\boldsymbol{W},$$

and $\boldsymbol{Z}_0 = \mathbf{0}$. If we assume that Δ is small enough (for daily data $\Delta = 1/252$, which is small enough), we substitute Δ for $\mathrm{d}t$. Consequently, we can write $\boldsymbol{\varepsilon} = \Delta^{-1/2}\mathrm{d}\boldsymbol{W}$ approximately has the $\mathcal{N}(\mathbf{0}, \mathbf{I})$ distribution. It follows from $\boldsymbol{Y}_t = \Delta^{1/2}\boldsymbol{Z}_t + \boldsymbol{Y}_0$ that

$$\mathrm{d}\boldsymbol{Z}_t \approx \Delta^{1/2}\boldsymbol{\mu}_{\mathbf{Y}}(\Delta^{1/2}\boldsymbol{Z}_t + \boldsymbol{Y}_0, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}.$$

We are interested in the distribution of $\boldsymbol{Z}_\Delta$, depending on $\boldsymbol{Z}_0$ for Δ small. We perform another approximation. The fact that $\boldsymbol{Z}_0 = \mathbf{0}$ implies $\boldsymbol{Z}_\Delta = \boldsymbol{Z}_\Delta - \boldsymbol{Z}_0 \approx \mathrm{d}\boldsymbol{Z}_0$, which results in the final approximation

$$\boldsymbol{Z}_\Delta \mid \boldsymbol{Y}_0 \sim \mathcal{N}(\Delta^{1/2}\boldsymbol{\mu}_{\mathbf{Y}}(\boldsymbol{Y}_0, \boldsymbol{\theta}), \mathbf{I}).$$

After a little degree of effort we obtain that for Δ small the density function of $\boldsymbol{X}_\Delta$ conditioned on $\boldsymbol{X}_0$ is approximately

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}, \Delta, \boldsymbol{\theta} \mid \boldsymbol{X}_0) &= \frac{\Delta^{-1/2}}{(2\pi)^{n/2}} \det[\boldsymbol{\sigma}^{-1}(\boldsymbol{x}, \boldsymbol{\theta})] \\ &\quad \times \exp\left\{ -\frac{1}{2\Delta} \left\| \boldsymbol{\gamma}(\boldsymbol{x}, \boldsymbol{\theta}) - \boldsymbol{\gamma}(\boldsymbol{X}_0, \boldsymbol{\theta}) - \Delta\boldsymbol{\mu}_{\mathbf{Y}}[\boldsymbol{\gamma}(\boldsymbol{X}_0), \boldsymbol{\theta}] \right\|^2 \right\}. \end{aligned}$$

Note also that

$$\boldsymbol{\mu}_{\mathbf{Y}}[\boldsymbol{\gamma}(\boldsymbol{X}_0), \boldsymbol{\theta}]\mathrm{d}t = \boldsymbol{\sigma}^{-1}(\boldsymbol{X}_0, \boldsymbol{\theta})\boldsymbol{\mu}(\boldsymbol{X}_0, \boldsymbol{\theta})\mathrm{d}t + \left.\begin{pmatrix} (\mathrm{d}\boldsymbol{X})'(\boldsymbol{\nabla}^2_{\boldsymbol{x}}\gamma_1(\boldsymbol{x}, \boldsymbol{\theta}))\mathrm{d}\boldsymbol{X} \\ \vdots \\ (\mathrm{d}\boldsymbol{X})'(\boldsymbol{\nabla}^2_{\boldsymbol{x}}\gamma_n(\boldsymbol{x}, \boldsymbol{\theta}))\mathrm{d}\boldsymbol{X} \end{pmatrix}\right|_{\boldsymbol{x}=\boldsymbol{X}_0},$$

where $(\mathrm{d}W_i)^2 = \mathrm{d}t$ and $(\mathrm{d}t)^2 = \mathrm{d}t\mathrm{d}W_i = 0$. Consequently, the density is invariant with respect to the addition of a constant vector to $\boldsymbol{\gamma}(\boldsymbol{x})$.

One might ask why we do not discretise the original model and use it as an approximation. The reason is that in the case of correlated increments of Wiener processes, the covariance matrix might be unstable and close to singular. For instance, in the model (2.1) the increment would have the normal distribution with the mean vector equal to drift multiplied by Δ and the covariance matrix

$$\begin{pmatrix} \sigma_{\mathrm{d}}^2 r_{\mathrm{d}}\Delta & \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}\sqrt{r_{\mathrm{d}}r_{\mathrm{u}}}\Delta \\ \rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}\sqrt{r_{\mathrm{d}}r_{\mathrm{u}}}\Delta & \sigma_{\mathrm{u}}^2 r_{\mathrm{u}}\Delta \end{pmatrix}.$$

Clearly, all the elements are almost of the same order, which might cause instability in the optimization (the increase in the objective maximum likelihood function in some directions might be very small or the optimization algorithm might try to make the matrix singular). One would have to make an orthogonal transformation (rotation) for each density in the log-likelihood function. On the other hand, in the presented approach the covariance matrix is diagonal, which implies higher stability.

In the following two subsections we exhibit approximate densities for the convergence model of Vasicek type (1.11) and the convergence model of CIR type (2.1).

3.6 Parameter estimation in the convergence model of Vasicek type

The constant diffusion matrix in the convergence model of Vasicek type simplifies all the computations. To shorten all the expressions we set $\boldsymbol{\theta} = (a, b, c, d, \sigma_{\mathrm{d}}, \sigma_{\mathrm{u}}, \rho)'$. We recall that in this model we have

$$\begin{aligned} \boldsymbol{\mu}(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) &= \begin{pmatrix} a + b(r_{\mathrm{u}} - r_{\mathrm{d}}) \\ c(d - r_{\mathrm{u}}) \end{pmatrix}, \\ \boldsymbol{\sigma}(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) &= \boldsymbol{\sigma}(\boldsymbol{\theta}) = \begin{pmatrix} \sqrt{1-\rho^2}\sigma_{\mathrm{d}} & \rho\sigma_{\mathrm{d}} \\ 0 & \sigma_{\mathrm{u}} \end{pmatrix}, \end{aligned}$$

where the diffusion matrix is rewritten in a form for independent increments of Wiener processes. Equation (3.11) implies that

$$\begin{aligned} \mathbf{J}_{(r_{\mathrm{d}}, r_{\mathrm{u}})}\boldsymbol{\gamma}(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) &= \begin{pmatrix} \partial_{r_{\mathrm{d}}}\gamma_1(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) & \partial_{r_{\mathrm{u}}}\gamma_1(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) \\ \partial_{r_{\mathrm{d}}}\gamma_2(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) & \partial_{r_{\mathrm{u}}}\gamma_2(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) \end{pmatrix} \\ &= \boldsymbol{\sigma}^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{1}{\sqrt{1-\rho^2}\sigma_{\mathrm{d}}} & -\frac{\rho}{\sqrt{1-\rho^2}\sigma_{\mathrm{u}}} \\ 0 & \frac{1}{\sigma_{\mathrm{u}}} \end{pmatrix}. \end{aligned}$$

Hence we obtain a system of partial differential equations, which gives us that

$$\boldsymbol{\gamma}(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) = \begin{pmatrix} \frac{r_{\mathrm{d}}}{\sqrt{1-\rho^2}\sigma_{\mathrm{d}}} - \frac{\rho r_{\mathrm{u}}}{\sqrt{1-\rho^2}\sigma_{\mathrm{u}}} \\ \frac{r_{\mathrm{u}}}{\sigma_{\mathrm{u}}} \end{pmatrix}.$$

The Hessian matrix is zero, which follows directly from the linearity of $\boldsymbol{\gamma}$. Therefore, $\boldsymbol{\mu}_{\mathbf{Y}}[\boldsymbol{\gamma}(r_{\mathrm{d}}^0, r_{\mathrm{u}}^0, \boldsymbol{\theta}), \boldsymbol{\theta}] = \boldsymbol{\sigma}^{-1}(\boldsymbol{\theta})\boldsymbol{\mu}(r_{\mathrm{d}}^0, r_{\mathrm{u}}^0, \boldsymbol{\theta})$. The determinant of matrix $\boldsymbol{\sigma}^{-1}(\boldsymbol{\theta})$ is

b	σ_{d}	ρ	λ_{d}	WMSE
5.83533	0.0482487	0.316187	−1.43618	0.0119852

Table 3.6: *The maximum likelihood estimates of the remaining parameters in the two-factor convergence model of Vasicek type and its weighted mean square error.*

$\det[\boldsymbol{\sigma}^{-1}(\boldsymbol{\theta})] = \frac{1}{\sqrt{1-\rho^2}\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}}$. So we are already able to employ the approximate density to estimate the parameters of the model.

Now we focus our attention on the estimation using the market data. It is a logical economic assumption that the evolution of the EMU's overnight interest rate is not influenced by the domestic interest rate. Therefore, the maximum likelihood estimates of parameters c, d, σ_{u} are obtained by the maximum likelihood estimator in the corresponding one-factor models for the EMU's data as well as λ_{d} is obtained by minimizing the WMSE in the one-factor model. We also set a equal to 0. The reason is that the period before adopting the Euro currency is too short to observe any minor divergence; hence, we assume that there is no minor divergence. In this way we reduced the problem of estimating 7 parameters to estimating 3 parameters, namely b, σ_{d}, ρ. We note that the main reason for employing a multivariate density is that we have to estimate the parameters in the SDE for the domestic rate simultaneously with estimating of ρ.

To estimate the remaining three parameters we applied a local maximization algorithm. The initial point was obtained by rounding the estimates in the one-factor Vasicek model for BRIBOR data, namely $b_0 = 29$ and $\sigma_{\mathrm{d}0} = 0.05$, and $\rho_0 = 0.5$. The resulting maximum likelihood estimates are presented in Table 3.6.

We estimate the market price of risk in the same manner as in the case of the one-factor models. We take the estimate of the EMU's market price of risk and minimize the weighted mean square error with respect to λ_{d}. That is, WMSE takes the form (3.10).

Table 3.6 displays the estimate of the market price of risk and the resulting WMSE. Figure 3.7 compares the estimated theoretical term structures with some selected true market term structure of interest rates.

3.7 Parameter estimation in the convergence model of CIR type

Compared with the convergence model of Vasicek type the only change is in the diffusion matrix (which is not constant)

$$\boldsymbol{\sigma}(r_{\mathrm{d}}, r_{\mathrm{u}}, \boldsymbol{\theta}) = \begin{pmatrix} \sqrt{1-\rho^2}\sigma_{\mathrm{d}}\sqrt{r_{\mathrm{d}}} & \rho\sigma_{\mathrm{d}}\sqrt{rd} \\ 0 & \sigma_{\mathrm{u}}\sqrt{ru} \end{pmatrix}.$$

b	σ_{d}	ρ	ν_{d}	WMSE
6.96169	0.285294	0.382321	−2.03	0.0112999

Table 3.7: *The maximum likelihood estimates of the remaining parameters in the two-factor convergence model of CIR type and its weighted mean square error.*

Therefore, transformation $\boldsymbol{\gamma}$ satisfies

$$\begin{aligned}\mathbf{J}_{(r_{\mathrm{d}},r_{\mathrm{u}})}\boldsymbol{\gamma}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) &= \begin{pmatrix}\partial_{r_{\mathrm{d}}}\gamma_1(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) & \partial_{r_{\mathrm{u}}}\gamma_1(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta})\\ \partial_{r_{\mathrm{d}}}\gamma_2(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) & \partial_{r_{\mathrm{u}}}\gamma_2(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta})\end{pmatrix}\\ &= \boldsymbol{\sigma}^{-1}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) = \begin{pmatrix}\frac{1}{\sqrt{1-\rho^2}\sigma_{\mathrm{d}}\sqrt{r_{\mathrm{d}}}} & -\frac{\rho}{\sqrt{1-\rho^2}\sigma_{\mathrm{u}}\sqrt{r_{\mathrm{u}}}}\\ 0 & \frac{1}{\sigma_{\mathrm{u}}\sqrt{r_{\mathrm{u}}}}\end{pmatrix},\end{aligned}$$

and we obtain a solution

$$\boldsymbol{\gamma}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) = \begin{pmatrix}\frac{2}{\sqrt{1-\rho^2}\sigma_{\mathrm{d}}}\sqrt{r_{\mathrm{d}}} - \frac{2\rho}{\sqrt{1-\rho^2}\sigma_{\mathrm{u}}}\sqrt{r_{\mathrm{u}}}\\ \frac{2}{\sigma_{\mathrm{u}}}\sqrt{r_{\mathrm{u}}}\end{pmatrix}.$$

Clearly, $\det[\boldsymbol{\sigma}^{-1}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta})] = \frac{1}{\sqrt{1-\rho^2}\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}\sqrt{r_{\mathrm{d}}r_{\mathrm{u}}}}$. To form an approximate density we have to compute $\boldsymbol{\mu}_{\mathbf{Y}}[\boldsymbol{\gamma}(r_{\mathrm{d}0},r_{\mathrm{u}0},\boldsymbol{\theta})]$, where $r_{\mathrm{d}0}$ and $r_{\mathrm{u}0}$ are the initial interest rates. It follows directly from $\mathbf{J}\boldsymbol{\gamma}$ that

$$\begin{aligned}\boldsymbol{\nabla}^2_{(r_{\mathrm{d}},r_{\mathrm{u}})}\gamma_1(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) &= \mathrm{diag}\left(\partial_{r_{\mathrm{d}}}(\sigma^{-1})_{11}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}),\partial_{r_{\mathrm{u}}}(\sigma^{-1})_{12}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta})\right),\\ \boldsymbol{\nabla}^2_{(r_{\mathrm{d}},r_{\mathrm{u}})}\gamma_2(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta}) &= \mathrm{diag}\left(0,\partial_{r_{\mathrm{u}}}(\sigma^{-1})_{22}(r_{\mathrm{d}},r_{\mathrm{u}},\boldsymbol{\theta})\right),\end{aligned}$$

where $\mathrm{diag}(\boldsymbol{x})$ represents a diagonal matrix with the vector $\boldsymbol{x}$ on its diagonal. Then we obtain that in this case

$$\begin{aligned}\boldsymbol{\mu}_{\mathbf{Y}}[\boldsymbol{\gamma}(r_{\mathrm{d}0},r_{\mathrm{u}0},\boldsymbol{\theta})] &= \boldsymbol{\sigma}^{-1}(r_{\mathrm{u}0},r_{\mathrm{d}0},\boldsymbol{\theta})\boldsymbol{\mu}(r_{\mathrm{u}0},r_{\mathrm{d}0},\boldsymbol{\theta})\\ &\quad+\left.\begin{pmatrix}\partial_{r_{\mathrm{d}}}(\sigma^{-1})_{11}(\sigma_{11}^2+\sigma_{12}^2)+\partial_{r_{\mathrm{u}}}(\sigma^{-1})_{12}\sigma_{22}^2\\ \partial_{r_{\mathrm{u}}}(\sigma^{-1})_{22}\sigma_{22}^2\end{pmatrix}\right|_{(r_{\mathrm{d}},r_{\mathrm{u}})=(r_{\mathrm{d}0},r_{\mathrm{u}0})}.\end{aligned}$$

The estimation is done in the same manner as in the case of the convergence model of Vasicek type. The constants c, d and σ_{u} were taken as the maximum likelihood estimates of the one-factor CIR model for the EMU's data, and a was set at 0. The remaining three parameters b, σ_{d} and ρ were estimated using local maximization with the initial values $b_0 = 25$, $\sigma_{\mathrm{d}0} = 0.3$ and $\rho = 0.5$ (rounded estimates in the one-factor CIR model for the Slovak data). The next step is to compute the constant in the market price of risk by minimizing the WMSE defined by equation (3.10).

Table 3.7 exhibits the resulting estimates for the two-factor convergence model of CIR type. We note that the process of minimization of ν_{d} is quite time-consuming in this model, and we can only use the zero-order information (i.e., no gradient information, etc.). The problem stems from the fact that the domestic term structure of interest rates is computed numerically. We were forced to figure out the optimal value of the constant in the market price of risk with a degree of precision to only two decimal places.

	One-factor models			Convergence models	
	Vasicek	CIR		Vasicek	CIR
κ	29.7431	25.486	a	0.0	0.0
θ	0.0224763	0.0220165	b	5.83533	6.96169
σ	0.0491624	0.288821	c	7.55337	7.75816
λ/ν	-10.8860	-43.5283	d	0.0219156	0.022182
			σ_{d}	0.0482487	0.285294
			σ_{u}	0.0214240	0.117200
			ρ	0.316187	0.382321
			$\lambda_{\mathrm{d}}/\nu_{\mathrm{d}}$	-1.43618	-2.03
			$\lambda_{\mathrm{u}}/\nu_{\mathrm{u}}$	-9.18920	-38.91972
WMSE	0.0141260	0.0112398		0.0119852	0.0112999

Table 3.8: *Calibration summary*

In Figure 3.7 we offer a visual comparison of the domestic term structures of interest rates for the convergence models of Vasicek and CIR type. Table 3.8 gives an overall view of the tested models for pricing the domestic discount bond and their degree of precision as measured by WMSE.

We note that we fix parameters, including the market price of risk, in the SDE for the EMU's rate in this estimation approach. Therefore, only one parameter (the domestic market price) left to fit the domestic term structures of interest rates. For example, if we do not fix the European market price of risk, we can exploit it to fit the market term structures. It is not correct, because the EMU's market price of risk is not influenced by the domestic market. Nevertheless, we tried this option. For the convergence model of Vasicek type we obtained $\lambda_{\mathrm{d}} = -6.99353$, $\lambda_{\mathrm{u}} = 1.29994$ and $\mathrm{WMSE} = 0.00832171$, and for the convergence model of CIR type we obtained $\nu_{\mathrm{d}} = -12.97$, $\nu_{\mathrm{u}} = -0.88$ and $\mathrm{WMSE} = 0.00799462$. The decrease in the mean square error is approximately 30% and 35%, respectively. We add that there is a possibility to use another estimation approach based on minimizing the WMSE through all parameters (see Ševčovič and Urbánová-Csajková (2005) for the one-factor CIR model). This approach should be more accurate, since the convergence model contains more parameters to fit the term structures than in the presented estimation approach. We leave this possibility for further research.

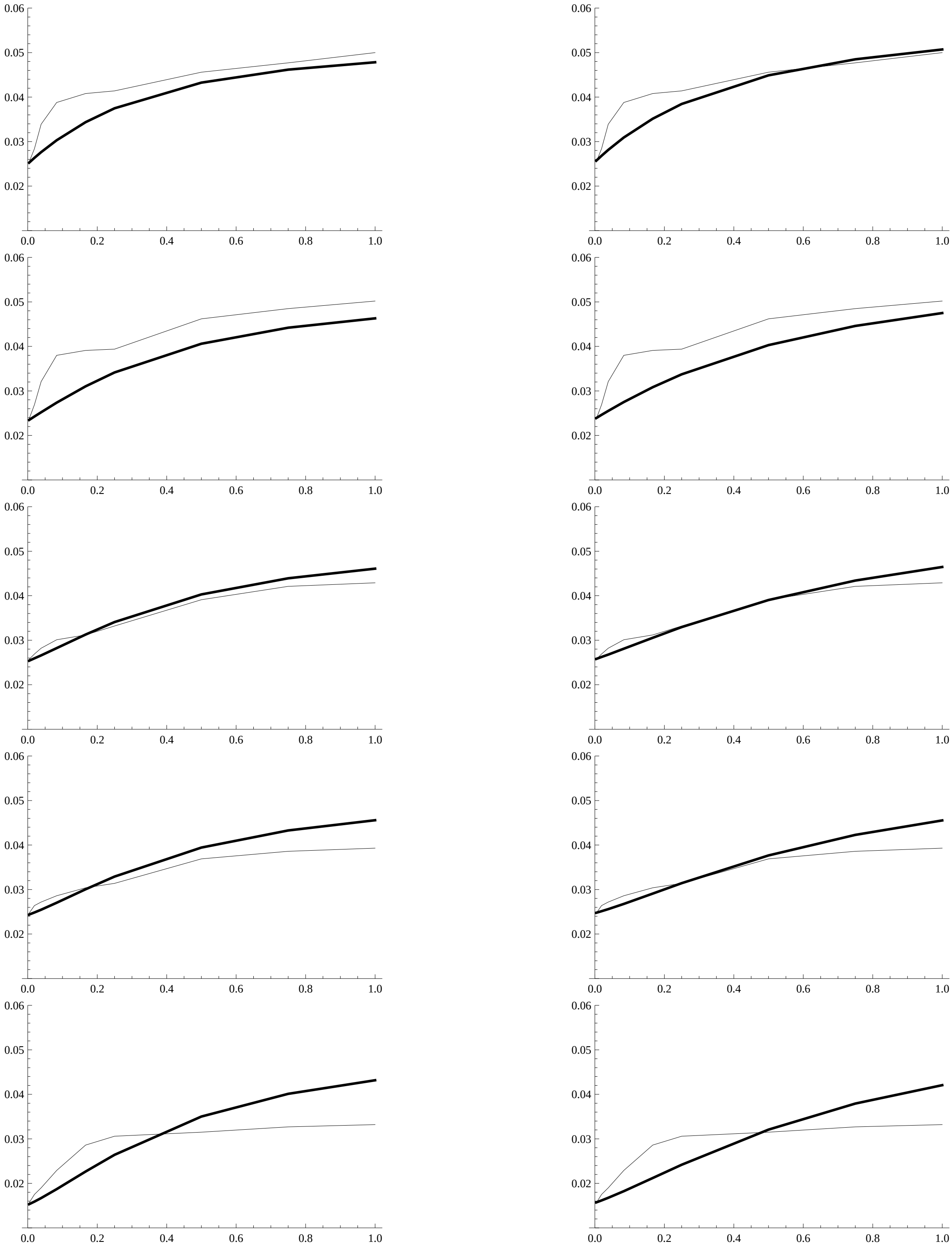

Figure 3.7: *Theoretical domestic term structures of interest rate (thick line) compared with the market term structures (thin line) in the two-factor convergence models of Vasicek type (left column) and CIR type (right column).*

Conclusion

In this thesis we study two-factor convergence term structure models of interest rates.

In the first part we focus our attention on the two-factor convergence model of Vasicek type proposed by Corzo and Schwartz (2000). We figure out a solution to the bond-pricing partial differential equation and compute the limit of the term structure of interest rates for the maturity going to infinity.

The second part deals with the proposed two-factor convergence model of Cox-Ingersoll-Ross (CIR) type, where the domestic and European overnight rates are governed by the Bessel square root process. We show that a separable solution exists only if the correlation between the increments of Wiener processes is zero; we derive its properties and compute the limit of the term structure of interest rates for the maturity going to infinity. In the other case we demonstrate that the separable solution for a zero correlation is a good approximation of the bond price in the case of nonzero correlation: we derive the order of approximation for the difference in logarithms of the bond prices with and without a correlation between the increments of Wiener processes.

Lastly, we compare the newly proposed two-factor convergence model of CIR type with the two-factor convergence model of Vasicek type as well as the one-factor Vasicek and CIR models. Inspired by papers by Aït-Sahalia (2002) and Aït-Sahalia (2008) we develop a simple and easy computable approximate density for multivariate diffusions. We employ the maximum likelihood estimator to find estimates for the parameters in the processes of the compared models. The market prices of risk are obtained by minimizing the weighted mean square error, which we use as an accuracy criterion. The results show that the mean square error of the proposed model is less than in the two-factor convergence model of Vasicek type. However, the one-factor CIR model gave the best result for this estimation approach.

This thesis offers two main possibilities for further research. Firstly, we can generalize the convergence model to the form of

$$\begin{aligned} \mathrm{d}r_{\mathrm{d}} &= [a + b(r_{\mathrm{u}} - r_{\mathrm{d}})]\mathrm{d}t + \sigma_{\mathrm{d}} r_{\mathrm{d}}^{\gamma_{\mathrm{d}}} \mathrm{d}W_{\mathrm{d}}, \\ \mathrm{d}r_{\mathrm{d}} &= c(d - r_{\mathrm{u}})]\mathrm{d}t + \sigma_{\mathrm{u}} r_{\mathrm{u}}^{\gamma_{\mathrm{u}}} \mathrm{d}W_{\mathrm{u}}, \\ \mathrm{Cov}[\mathrm{d}W_{\mathrm{d}}, \mathrm{d}W_{\mathrm{u}}] &= \rho \mathrm{d}t. \end{aligned}$$

Compared with the previously mentioned convergence models, in this case we expect non-existence of any separable solution. We note that there is no solution even for the one-factor CKLS model (Chan et al. (1992)). Therefore, an analytic approximate solution is needed (see, e.g., Choi and Wirjanto (2007) or Stehlíková and Ševčovič (2009) for the case of the one-factor CKLS model).

The second problem is to improve the calibration methodology. In the first step we suggest to apply the two-phase minmax method (see Ševčovič and Urbánová-Csajková (2005)) to estimate the model for pricing the EMU's bond, and then to minimize the weighted mean square error through the remaining parameters.

Résumé (in Slovak)

V tejto práci sa zaoberáme modelmi časovej štruktúry úrokových mier a oceňovaním dlhopisov. Jadro tvorí štúdium dvoch tzv. konvergenčných modelov úrokovej miery, ktorých vznik podnietil fakt, že medzi vývojom úroku Európskej Menovej Únie (EMÚ) a vývojom úroku v krajine vstupujúcej do EMÚ existuje korelácia. Na obrázku 1.5 na strane 10 je znázornený vývoj slovenskej a európskej jednodňovej úrokovej miery, čo potvrdzuje existenciu korelácie. Na modelovanie vývoja úrokových mier sme využili stochastické diferenciálne rovnice (SDR), ktoré spolu s Itōvou lemou, zaisťovaním dlhopisov a vylúčením arbitráže umožňujú previesť problém oceňovania dlhopisov na riešenie parciálnej diferenciálnej rovnice (PDR) (čitateľa odkazujeme na diela Øksendal (2000) a Kwok (1998)).

V prvej časti sa zaoberáme analýzou vlastností prvého konvergenčného modelu, ktorý navrhli Corzo and Schwartz (2000), pričom na jeho formuláciu využili Vašíčkov model (preto ho budeme ďalej nazývať konvergenčný model Vašíčkovho typu):

$$\begin{aligned}
\mathrm{d}r_\mathrm{d} &= [a + b(r_\mathrm{u} - r_\mathrm{d})]\mathrm{d}t + \sigma_\mathrm{d}\mathrm{d}W_\mathrm{d},\\
\mathrm{d}r_\mathrm{u} &= c(d - r_\mathrm{u})\mathrm{d}t + \sigma_\mathrm{u}\mathrm{d}W_\mathrm{u},\\
\mathrm{Cov}[\mathrm{d}W_\mathrm{d}, \mathrm{d}W_\mathrm{u}] &= \rho\mathrm{d}t,
\end{aligned}$$

kde r_d je domáca úroková miera, r_u je európska úroková miera, b, c, σ_d a σ_u sú kladné konštanty, a a d sú nezáporné konštanty, a $\rho \in (0, 1)$ je korelácia medzi prírastkami $\mathrm{d}W_\mathrm{d}$ a $\mathrm{d}W_\mathrm{u}$ Wienerových procesov. Príslušná oceňovacia PDR je

$$\begin{aligned}
-\partial_\tau P + [a + b(r_\mathrm{u} - r_\mathrm{d}) - \lambda_\mathrm{d}\sigma_\mathrm{d}]\partial_{r_\mathrm{d}} P + [c(d - r_\mathrm{u}) - \lambda_\mathrm{u}\sigma_\mathrm{u}]\partial_{r_\mathrm{u}} P + \\
\tfrac{1}{2}\sigma_\mathrm{d}^2\partial^2_{r_\mathrm{d}} P + \tfrac{1}{2}\sigma_\mathrm{u}^2\partial^2_{r_\mathrm{u}} P + \rho\sigma_\mathrm{d}\sigma_\mathrm{u}\partial^2_{r_\mathrm{d}r_\mathrm{u}} P - r_\mathrm{d}P &= 0,\\
P(r_\mathrm{d}, r_\mathrm{u}, 0) &= 1,
\end{aligned}$$

pričom τ predstavuje čas do maturity. Ak uvažujeme riešenie v tvare $P(r_\mathrm{d}, r_\mathrm{u}, \tau) = \exp\{A(\tau) - D(\tau)r_\mathrm{d} - U(\tau)r_\mathrm{u}\}$, potom riešenie uvedenej PDR prevedieme na riešenie systému obyčajných diferenciálnych rovníc (ODR) (1.19)–(1.21) (strana 12). Uvedený systém sme napokon vyriešili, nakoľko v pôvodnom článku bolo uvedené nesprávne riešenie. Pôvodné výsledky z článku Corzo and Schwartz (2000) sme rozšírili o vlastný prínos, ktorý sme sformulovali do uvedeného tvrdenia:

Tvrdenie 1. *Limita domácej časovej štruktúry úrokových mier v konvergenčnom modeli Vašíčkovho typu je*

$$\lim_{\tau\to\infty} R(r_{\rm d}, r_{\rm u}, \tau) = \frac{a}{b} + d - \frac{c^2\sigma_{\rm d} + b^2\sigma_{\rm u}(2c\lambda_{\rm u} + \sigma_{\rm u}) + 2bc\sigma_{\rm d}(c\lambda_{\rm d} + \rho\sigma_{\rm u})}{2b^2c^2}.$$

V druhej časti práce sa zaoberáme novým, nami navrhnutým, konvergenčným modelom, na ktorého formuláciu sme využili Cox-Ingersoll-Rossov model (preto ho nazývame konvergenčný model typu CIR):

$$\begin{aligned} {\rm d}r_{\rm d} &= [a + b(r_{\rm u} - r_{\rm d})]{\rm d}t + \sigma_{\rm d}\sqrt{r_{\rm d}}{\rm d}W_{\rm d}, \\ {\rm d}r_{\rm u} &= c(d - r_{\rm u}){\rm d}t + \sigma_{\rm u}\sqrt{r_{\rm u}}{\rm d}W_{\rm u}, \\ {\rm Cov}[{\rm d}W_{\rm d}, {\rm d}W_{\rm u}] &= \rho{\rm d}t. \end{aligned}$$

Na rozdiel od konvergenčného modelu Vašíčkovho typu, nami navrhnutý model nepripúšťa záporné hodnoty úrokových mier. Opäť sme hľadali riešenie v tvare $P(r_{\rm d}, r_{\rm u}, \tau) = \exp\{A(\tau) - D(\tau)r_{\rm d} - U(\tau)r_{\rm u}\}$. V prípade $\rho = 0$ sme previedli oceňovaciu PDR na systém ODR (viď systém (2.5)–(2.7) na strane 19), pre ktorý sa nám nepodarilo nájsť explixitné riešenie, dá sa však riešiť numericky. Pre prípad $\rho = 0$ sme ďalšie vlastnosti funkcií A, D a U, ako aj domácej časovej štruktúry úrokových mier, sformulovali do nasledovných tvrdení:

Lema 2. *Nech $A(\tau)$, $D(\tau)$ a $U(\tau)$ riešia systém ODR* (2.5)–(2.7). *Potom:* ***i)*** $D(\tau) > 0$ *je monotónna, rastúca a* $\lim_{\tau\to\infty} D(\tau) = D_\ominus$ *(viď* (2.9) *na strane 19),* ***ii)*** $U(\tau) > 0$ *je monotónna, rastúca a ohraničená, a* ***iii)*** *ak* $a \geq 0$ *potom* $A(\tau) < 0$*; pre všetky* $\tau > 0$.

Dôsledok 3. *Limita $U(\tau)$ pre $\tau \to \infty$ je*

$$\widehat{U} = \lim_{\tau\to\infty} U(\tau) = \frac{(c + \sigma_{\rm u}\nu_{\rm u}) - \sqrt{(c + \sigma_{\rm u}\nu_{\rm u})^2 + 2b\sigma_{\rm u}^2 D_\ominus}}{-\sigma_{\rm u}^2}.$$

Tvrdenie 4. *Limita domácej časovej štruktúry úrokových mier v konvergenčnom modeli typu CIR je*

$$\lim_{\tau\to\infty} R(r_{\rm d}, r_{\rm u}, \tau) = aD_\ominus + cd\frac{(c + \nu_{\rm u}\sigma_{\rm u}) - \sqrt{(c + \nu_{\rm u}\sigma_{\rm u})^2 + 2bD_\ominus\sigma_{\rm u}^2}}{-\sigma_{\rm u}^2}.$$

Pre prípad $\rho \neq 0$ sme ukázali, že sa oceňovacia PDR rozpadne na systém nekonzistentných rovníc. Preto sme sa rozhodli aproximovať riešenie PDR pre $\rho \neq 0$ riešením PDR pre $\rho = 0$. Chyba tejto aproximácie je sformulovaná v nasledujúcej vete:

Veta 5. *Nech* $P_{\rm CIR}(r_{\rm d}, r_{\rm u}, \tau; \rho)$ *je riešenie oceňovacej PDR* (2.3) *pre konvergenčný model typu CIR. Potom*

$$\ln[P_{\rm CIR}(r_{\rm d}, r_{\rm u}, \tau; 0)] - \ln[P_{\rm CIR}(r_{\rm d}, r_{\rm u}, \tau; \rho)] = c_3(r_{\rm d}, r_{\rm u}; \rho)\tau^3 + {\rm o}(\tau^3),$$

kde koeficient c_3 nie je identicky rovný nule.

Pre porovnanie:

Tvrdenie 6. *Nech* $P_{\mathrm{Vas}}(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau; \rho)$ *je riešenie oceňovacej PDR* (1.17) *pre konvergenčný model Vašíčkovho typu. Potom*

$$\ln[P_{\mathrm{Vas}}(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau; 0)] - \ln[P_{\mathrm{Vas}}(r_{\mathrm{d}}, r_{\mathrm{u}}, \tau; \rho)] = -\frac{1}{8} b\rho\sigma_{\mathrm{d}}\sigma_{\mathrm{u}}\, \tau^4 + \mathrm{o}(\tau^4).$$

Pre odhady parametrov z článku Corzo and Schwartz (2000) sa tento rozdiel prejaví na jednej stotine percenta až pre dlhopisy s maturitou 20 rokov (viď tabuľka 2.1 na strane 23). Reálne sa však diskontné dlhopisy vydávajú na krátke obdobia, preto môžeme túto aproximáciu považovať za dobrú.

Posledná časť práce sa zaoberá kalibráciou a následným porovnaním konvergenčných modelov medzi sebou, ako aj s vybranými jednofaktorovými modelmi. Modely sme sa rozhodli kalibrovať nasledovne: najprv z trhových dát odhadneme parametre procesu pre okamžité úrokové miery, a následne minimalizáciou váženej strednej kvadratickej chyby cez konštanty obsiahnuté v trhovej cene rizika nafitujeme výnosové krivky na trhové výnosové krivky. Parametre procesu odhadujeme metódou maximálnej vierohodnosti. Na to však potrebujeme poznať hustotu príslušného procesu, čo nie je možné vždy zabezpečiť. Nájsť hustotu nejakého procesu

$$\mathrm{d}\boldsymbol{X} = \boldsymbol{\mu}(\boldsymbol{X}, t, \boldsymbol{\theta})dt + \boldsymbol{\sigma}(\boldsymbol{X}, t, \boldsymbol{\theta})\mathrm{d}\boldsymbol{W},$$

hľadaním presného riešenia v uzavretom tvare tzv. Fokker-Planckovej rovnice nie je vždy možné (skôr ojedinelé). Inšpirovaní článkami Aït-Sahalia (2002) a Aït-Sahalia (2008) sme odvodili približnú hustotu procesu $\boldsymbol{X}_t$

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x}, \Delta, \boldsymbol{\theta} \mid \boldsymbol{X}_0) \quad &= \quad \frac{\Delta^{-1/2}}{(2\pi)^{n/2}} \det[\boldsymbol{\sigma}^{-1}(\boldsymbol{x}, \boldsymbol{\theta})] \\ &\quad \times \exp\left\{ -\frac{1}{2\Delta} \left\| \boldsymbol{\gamma}(\boldsymbol{x}, \boldsymbol{\theta}) - \boldsymbol{\gamma}(\boldsymbol{X}_0, \boldsymbol{\theta}) - \Delta\boldsymbol{\mu}_{\mathbf{Y}}[\boldsymbol{\gamma}(\boldsymbol{X}_0), \boldsymbol{\theta}] \right\|^2 \right\}, \end{aligned}$$

kde $\boldsymbol{\gamma}$ je také bijektívne zobrazenie, ktoré rieši $\mathbf{J}_{\boldsymbol{x}}\boldsymbol{\gamma}(\boldsymbol{x}, \boldsymbol{\theta}) = \boldsymbol{\sigma}^{-1}(\boldsymbol{x}, \boldsymbol{\theta})$ (hovoríme, že proces je reducibilný), $\boldsymbol{\mu}_{\mathbf{Y}}$ je drift procesu $\boldsymbol{Y} = \boldsymbol{\gamma}(\boldsymbol{X}, \boldsymbol{\theta})$ a Δ je časový rozdiel medzi dvoma po sebe idúcimi pozorovaniami procesu $\boldsymbol{X}_t$. Parametre popisujúce vývoj európskej úrokovej miery boli odhadované z príslušného jednofaktorového modelu, a následne boli odhadované parametre pre vývoj domáceho úroku. Odhady v jednotlivých modeloch sú zhrnuté v tabuľke 3.8 na strane 45. Na základe strednej kvadratickej chyby modelov konštatujeme, že chyba nami navrhnutého konvergenčného modelu je nižšia ako pôvodného konvergenčného modelu, no nie je nižšia ako chyba jednofaktorového CIR modelu. Poznamenávame, že v prípade uvoľnenia európskej trhovej ceny rizika (čo však nie je korektné) sa chyba konvergenčných modelov znížila o rádovo 30% (stredná kvadratická chyba bola výrazne nižšia ako v prípade jednofaktorových modelov). Preto sa tu otvára priestor pre iné metódy kalibrácie konvergenčných modelov.

Bibliography

Aït-Sahalia, Y.: 1999, Transition Densities for Interest Rate and Other Nonlinear Diffusions, *Journal of Finance* **54**(4), 1361–1395.

Aït-Sahalia, Y.: 2002, Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach, *Econometrica* **70**(1), 223–262.

Aït-Sahalia, Y.: 2008, Closed-form Likelihood Expansions for Multivariate Diffusions, *The Annals of Statistics* **36**(2), 906–937.

Anderson, T. G. and Lund, J.: 1996, The Short Rate Diffusion Revisited: a Tour Guided by the Efficient Method of Moments, *Working paper of J. L. Kellogg Graduate School of Management, Northwestern University* .

Balduzzi, P., Das, S. R. and Foresi, S.: 1998, The central tendency: A second factor in bond yields, *Review of Economics and Statistics* **80**, 62–72.

Chan, K. C., Karolyi, G. A., Longstaff, F. and Sanders, A.: 1992, The volatility of short-term interest rates: an empirical comparison of alternative models of the term structure of interest rates, *Journal of Finance* **47**, 1209–1227.

Choi, Y. and Wirjanto, T. S.: 2007, An analytic approximation formula for pricing zero-coupon bonds, *Finance Research Letters* **4**, 116–126.

Corzo, T. and Schwartz, E. S.: 2000, Convergence within the European Union: Evidence from Interest Rates, *Economic Notes* **29**(2), 243–268.

Cox, J. C., Ingersoll, Jr., J. E. and Ross, S. A.: 1985, A theory of the term structure of interest rates, *Econometrica* pp. 385–407.

Evans, L. C.: 1998, *Partial Differential Equations*, American Mathematical Society.

Fong, H. G. and Vasicek, O. A.: 1991, Fixed–income Volatility Management, *Journal of Portfolio Management* **17**(4), 41–46.

Gallant, A. R. and Tauchen, G.: 1996, Which Moments to Match?, *Econometric Theory* **12**, 657–681.

Gouriéroux, C., Monfort, A. and Renault, E.: 1993, Indirect Inference, *Journal of Applied Econometrics* **8**, S85–S118.

Hansen, L. P. and Scheinkman, J. A.: 1995, Back to the Future: Generating Moment Implications for Continuous Time Markov Processes, *Econometrica* **63**, 767–804.

Itō, K.: 1951, On a Formula Concerning Stochastic Differentials, *Nagoya Math. Journal* **3**, 55–65.

Kessler, M. and Sorensen, M.: 1999, Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion, *Bernoulli* **5**, 299–314.

Kwok, Y. K.: 1998, *Mathematical Models of Financial Derivatives*, 1 edn, Springer.

Kwok, Y. K.: 2008, *Mathematical Models of Financial Derivatives*, 2 edn, Springer.

Øksendal, B.: 2000, *Stochastic Differential Equations*, 5 edn, Springer.

Stehlíková, B. and Ševčovič, D.: 2005, On a Volatility Averaging in a Two-Factor Interest Rate Model, *Proceedings of Algoritmy 2005*, pp. 325–333.

Stehlíková, B. and Ševčovič, D.: 2009, Approximate Formulae for Pricing Zero-Coupon Bonds and their Asymptotic Analysis, *International Journal of Numerical Analysis and Modeling* **6**, 274–283.

Vasicek, O. A.: 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**(2), 177–188.

Ševčovič, D.: 2008, *Partial Differential Equations and their Applications*, Iris.

Ševčovič, D., Stehlíková, B. and Mikula, K.: 2009, *Analytical and Numerical Methods for Pricing Financial Derivatives*, 1 edn, STU.

Ševčovič, D. and Urbánová-Csajková, A.: 2005, On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model, *CEJOR* **13**, 169–188.