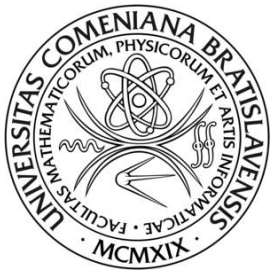
COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS



FONG-VASICEK MODEL OF INTEREST RATES WITH STOCHASTIC VOLATILITY

DIPLOMA THESIS

2014

Martin ŠIMO

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

FONG-VASICEK MODEL OF INTEREST RATES WITH STOCHASTIC VOLATILITY

DIPLOMA THESIS

Study programme:	Economic and Financial Mathematics
Field of Study:	9.1.9. Applied Mathematics
Department:	FMFI.KAMŠ - Department of Applied Mathematics and Statistics
Supervisor:	RNDr. Mgr. Beáta Stehlíková, PhD.
Evidence number:	1ec9dfab-8b52-419f-a609-6f9ac4e12ee7

Bratislava 2014 **Martin ŠIMO**

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Bc. Martin Šimo
Study programme:	Economic and Financial Mathematics (Single degree study, master II. deg., full time form)
Field of Study:	9.1.9. Applied Mathematics
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Fong-Vasicek model of interest rates with stochastic volatility

Aim: Fong-Vasicek model of interest rates contains volatility which is not observable on the market. In Master thesis (Selečéniová 2012), an approximation of bond prices has been derived. It contains a smaller number of parmeters compared to the original model and it is also independent of the current level of volatility. The aim of the thesis is to propose and perform a calibration procedure based on this approximation.

Supervisor:	RNDr. Mgr. Beáta Stehlíková, PhD.
Department:	FMFI.KAMŠ - Department of Applied Mathematics and Statistics
Head of department:	prof. RNDr. Daniel Ševčovič, CSc.

Assigned: 25.01.2013

Approved: 04.02.2013 prof. RNDr. Daniel Ševčovič, CSc.
Guarantor of Study Programme

.....
Student

.....
Supervisor

I declare this thesis was written on my own, with the only help provided by my supervisor and the referenced liturature.

Bratislava, 28.4.2014 **Bc. Martin Šimo**

I would like to thank my supervisor RNDr. Mgr. Beáta Stehlíková, PhD. for prvoviding guidance in research and having patience with me.

Abstrakt

ŠIMO, Martin: Fong-Vašíčkov model úrokových mier so stochastickou volatilitou [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; vedúci práce: RNDr. Mgr. Beáta Stehlíková, PhD., Bratislava, 2014, 54s.

Táto práca sa venuje modelom okamžitej úrokovej miery. V práci predstavujeme ich základné myšlienky a skúmame dva takéto modely. Menovite skúmame jednofaktorový Vašíčkov model a dvojfaktorový Fong-Vašíčkov model so stochastickou volatilitou. Predstavujeme aj myšlienku rýchlej škály volatility a na nej založenú asymptotickú aproximáciu Fong-Vašíkovho modelu. V práci definujeme algoritmus, ktorým odhadneme parametre tejto aproximácie z výnosových kriviek. Tento algoritmus potom testujeme na náhodne generovaných aj reálnych dátach a skúmame jeho správnosť a užitočnosť.

Kľúčové slová: Oceňovanie dlhopisov, Modely okamžitej úrokovej miery, Vašíčkov model, Fong-Vašíčkov model, Stochastická volatilita, Rýchla škála volatility, Kalibrácia modelov okamžitej úrokovej miery

Abstract

ŠIMO, Martin: Fong-Vasicek model of interest rates with stochastic volatility [Diploma Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: RNDr. Mgr. Beáta Stehlíková, PhD., Bratislava, 2014, 54p.

This thesis focuses on the study of short rate models. It describes their basic notions and defines two such models, namely the one factor Vasicek model and the two factor Fong-Vasicek model with stochastic volatility. An idea of fast scaling volatility is introduced and an asymptotic approximation of Fong-Vasicek model that is based on this notion is described. We devise an algorithm to estimate the parameters of this approximation from yield curve data. This algorithm is then then tested on both randomly generated and real world data and its validity and usefulness is assessed.

Keywords: Bond pricing, Short rate model, Vasicek model, Fong-Vasicek model, Stochastic volatility, Fast scaling volatility, Short rate model calibration

Contents

Introduction

This work focuses on short rate models. These models became increasingly important with the rapid growth of the interest rate derivatives markets. The increased popularity of these markets called for a way of pricing these derivatives. The short rate models were defined for this purpose. They build upon the often used framework for pricing continous time derivatives defined by Black and Scholes in [3].

One of the earliest short rate models was described by Vasicek in [27]. The one factor Vasicek model is based on a stochastic differential equation that governs the stochastic behaviour of the short rate. We describe this model in the first section of this work. Later on it became apparent that a one factor model cannot satisfyingly describe the behaviour of the markets. To allow for more information to be captured by a model two factor models were defined. In this thesis we work with one such model, namely the Fong-Vasicek model with stochastic volatility first described in [13]. This model takes the volatility, which is a constant in Vasicek model and assumes that it is also a stochastic process. Some motivation for picking volatility as the second factor along with the description of this model can also be found in the first section.

We introduce an asymptotic approximation for bond pricing function of the Fong-Vasicek model, described in [9] and [20]. This approximation is based on the idea of fast scaling volatility, which assumes that the short rate process and the volatility process operate on different time scales. These notions along with the implications and results of this idea are contained in the second section.

We also define a data generating process that allows us to generate observations of short rate and volatility that follow the Fong-Vasicek model. This idea is based on a straightforward discretization of the stochastic differential equations defining these models. This data generating process along with the parameters we use is described in the third section.

The fourth section is the most important section of this work. In it we define an algorithm that estimates the parameters of the Vasicek model and the asymptotic approximation of Fong-Vasicek model. The idea behind the algorithm is based on

fitting the observed yield curves. The section also contains some general ideas about the estimation of short rate models.

The last section contains the results of the estimation algorithm defined in fourth section. The models are first fitted onto generated data to assess whether the estimation algorithm is valid. This is done by fitting models that were used to generate the data in the first place, where we expect an almost exact fit. The second part of the section details the results obtained by employing the estimation algorithms on real world data about German yield curves.

1 Short rate models

In this section we define the basic concepts and terms we will use throughout the work, the main resources for this part are textbooks [18] and [22]. We then introduce the notion of short rate models and the motivations for their study. We also formally introduce the two short rate models that we use in this work.

1.1 Basic concepts

A bond represents a contract between two parties, the issuer and the holder of the bond. The issuer takes up an obligation to pay an amount F, called the face value, to the holder of the bond at a set time in the future, called the maturity of a bond. He is also obliged to pay periodic amounts C, called the coupon, at predetermined dates leading up to the maturity. A bond which pays no coupons is called a zero-coupon bond. The ownership of the bond is in general transferable giving rise to a market of bonds.

In this thesis we work with discount bonds, these are zero-coupon bonds with face value equal to 1. On the market for bonds a price of a discount bond, with various maturities can be observed. Each of these prices $P(T_0, T)$ implies an interest rate $R(T_0, T)$ and if we consider continuously compounded interest, the relationship between the price of the bond and the interest rate it implies is captured by the formula

$$P(T_0, T) = e^{-R(T_0,T)(T-T_0)}. \tag{1.1}$$

The individual interest rates $R(T_0, T)$ are called the spot rates of a bond and they depend on the difference $\tau = T_0 - T$, called time to maturity. The relationship between time to maturity τ and the interest rate of a discount bond is called the zero-coupon yield curve or the term structure of interest rates. For the remained of this thesis we call zero-coupon yield curves simply yield curves.

The short rate r_t is a term used for the beginning of the yield curve $R(T, T)$. We know from equation (1.1), that

$$R(T_0, T) = -\frac{1}{T_0 - T} \ln P(T_0, T). \tag{1.2}$$

Considering a bond with a time to maturity τ we arrive at

$$R(t, t+\tau) = -\frac{1}{\tau} \ln P(t, t+\tau) = -\frac{1}{\tau} \left(\ln P(t, t+\tau) - \ln P(t, t) \right),$$

where we used the fact that $\ln P(t, t) = 0$ since $P(t, t) = 1$ must obviously be true. If we now consider a limit for $\tau \to 0$ we arrive at

$$r_t = R(t, t) = \lim_{\tau \to 0} -\frac{1}{\tau} \left(\ln P(t, t+\tau) - \ln P(t, t) \right) = -\frac{\partial}{\partial T} \ln P(t, t). \tag{1.3}$$

Knowledge of the short rate does not automatically imply the knowledge of the whole yield curve, as the short rate is only its beginning. However there are models which assume this to be the case. After this property such models are called short rate models.

1.2 Short rate models

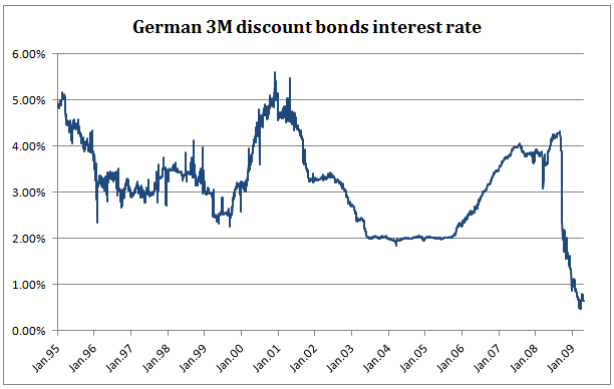


Figure 1.1: The interest rate implied by German bonds with maturity 3 months.

To illustrate the behaviour of the short rate we refer to the Figure 1.1, which shows the development of the interest rate implied by German bonds with maturity of 3 months. This is the shortest maturity we use in this work and we consider it an appropriate approximation of the short rate, similar approach was adopted in [1]. We can clearly see the erratic behaviour of the short rate. In short rate models this fact is dealt with by assuming that the short rate r_t follows a stochastic process defined by a stochastic differential equation. Based on the specifics of the equation the models are then divided.

These models could come into being because of the surge in interest caused by Black and Scholes and their, to this day, very popular work [3]. Their approach allowed for simplicity and parsimony in defining contionuous time models. These continous time models were extremely appealing in theory, because they are much easier to work with than discrete time models. This attractive feature leads to the fact that many solutions of models based on this framework are analytical in nature, meaning that the results of the models can be captured by formulas.

The first and simplest class of short rate models is the class of one factor models. In one factor models the short rate process r_t is said to be a solution of a stochastic differential equation of the form

$$dr = \mu(t, r)dt + \sigma(t, r)dW. \tag{1.4}$$

The process consists of two parts. The deterministic drift factor is captured by the term $\mu(t, r)dt$. This term is responsible for the trends in the behaviour of the short rate. The second term $\sigma(t, r)dW$, where dW is an increment of a Wiener process W, is the volatility factor. This factor is responsible for the random deviations of the short rate around the direction defined by drift. For a detailed list of different one factor short rate models and their properties we refer the reader to the works [5] and [15].

The second class of models we use touch on in this work is the class of two factor models. These models postulate that the short rate $r = r(x, y)$ is a function of two factors x, y. Both x and y are governed by a stochastic differential equation. Their joint behaviour is captured by the system

$$dx = \mu_x(x, y)dt + \sigma_x(x, y)dW_1, \tag{1.5}$$

$$dy = \mu_y(x, y)dt + \sigma_y(x, y)dW_2, \tag{1.6}$$

where the correlation of increments dW_1 and dW_2 of the Wiener processes W_1 and W_2 is a constant ρ, i.e. $\mathrm{E}(dW_1 dW_2) = \rho dt$. The function joining the factors into the short rate can be of different forms. For example the short rate r can be a sum of x and y. Another example is a model that postulates that a national short rate process is in some way affected by a short rate process of a different country, e.g. the stochastic

short rate in France is probably affected by the again stochastic short rate for EU bonds. The last example, the type of two factor model we use in this work, is a model in which one of the parameters of the one factor model is assumed to be stochastic in nature, e.g. a one factor model where we consider the volatility to be stochastic, thus making it the second factor.

1.3 Vasicek Model

First described by Oldrich Vasicek in [27], the Vasicek model is one of the earliest models of short rate. It is a one factor short rate model implying that the yield curve R depends solely on the short rate. The short rate is a random process defined by a stochastic differential equation

$$dr = \kappa(\theta - r)dt + \sigma dW, \tag{1.7}$$

where dW are increments in a Wiener process W. Being one of the earliest models it is very simple, with all parameters deterministic. The drift term in this model has a mean-reverting property, where θ is the long term average short rate and κ the speed of reversion towards it. The processes with the mean-reverting property are also known as Ornstein-Uhlbeck type processes. Over time this model has been found to be insufficient to capture the behaviour observed in practice, where the assumption of one factor has proven to be too strong, as evidenced in works [10] and [19]. It also has theoretical problems, namely it can lead to negative interest rates with non-trivial probability, which is something that should not be possible, at least not for nominal interest rates. The model's simplicity however means that we can easily use it as a reference model to be compared with other more complex models.

The pricing function $P(\tau, r)$ for a discount bond, depending on time to maturity τ and the value of short rate r, is a solution to a stochastic differential equation

$$-\frac{\partial P}{\partial \tau} + (\mu(t,r) - \lambda(t,r)\sigma(t,r))\frac{\partial P}{\partial r} + \frac{\sigma^2(t,r)}{2}\frac{\partial^2 P}{\partial r^2} - rP = 0, \tag{1.8}$$

with the initial condition $P(0, r) = 1$. For the proof of the previous statement we refer the reader to [24]. The proof uses the no-arbitrage approach and Ito lemma to arrive

at this result. In the proof $\lambda(t,r)$ arises as function defining the market value of risk. This function provides an expected rise of the bond return for the unit rise of risk, for more details see [15].

If the short rate follows the process defined by (1.7) and we assume a constant market value of risk $\lambda(t,r) = \lambda$ the solution to (1.8) is of the form

$$P(\tau, r) = A(\tau)e^{-B(\tau)r}. \tag{1.9}$$

The particular functions $A(\tau)$ and $B(\tau)$ are found by substituting a solution of this form into (1.8). This yields a system of ODE

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

with initial conditions $A(0) = 1$ and $B(0) = 0$. The solution to this system is given by

$$A(\tau) = \exp\left[(-\theta + \frac{\sigma^2}{2\kappa^2} + \frac{\sigma\lambda}{\kappa})(-\frac{1}{\kappa}(1 - e^{-\kappa\tau}) + \tau) - \frac{\sigma^2}{4\kappa^3}(1 - e^{-\kappa\tau})^2\right], \tag{1.10}$$

$$B(\tau) = -\frac{(1 - e^{-\kappa\tau})}{\kappa}. \tag{1.11}$$

To give an idea about how the short rate process defined by Vasicek model might look in time we provide Figure 1.2. In this figure we can see the property of mean-reversion typical for Vasicek model.

1.4 Stochastic Volatility

One of the assumptions of the Black Scholes framework defined in [3] was the assumption of constant volatility. This was an assumption even the authors considered violated in practice, noting this fact in [2]. They urge for more work to be done in the field of predicting volatility from available data. The fact that volatility is not constant has also been evidenced in other empirical work, see e.g. [16]. Another example of changing volatility is the volatility smile. An observation arising from prices of some financial derivatives, that the relashionship between implied volatility and strike price has a shape of a smile, for more details see [12].

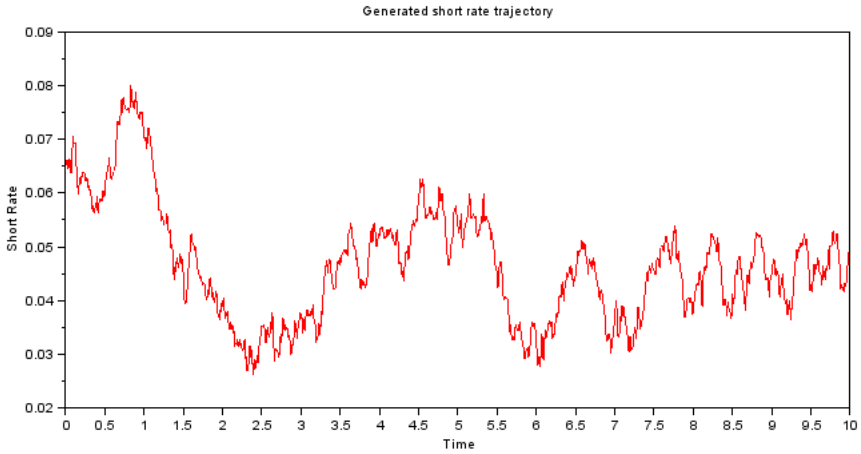


Figure 1.2: A simulated development of short rate r that follows Vasicek model.

For these reasons many models have adopted volatility as an additional stochastic factor. In short rate models these include a two factor Cox-Ingersoll-Ross model, see [16], stochastic Vasicek model or Brenner model [4]. Longstaff and Schwarts in [16] also argue that stochastic volatility allows for more shapes of yield curves, such that seem to occur in practical observations.

Calculating the estimations of volatility in short rate data a number of stylized facts have been observed over and over again. One of the most often found patters is the so called volatility clustering. This pattern, described as early as 1963 in [17], can be desribed by changing periods of high volatility alternating with periods with low probability. Thus the volatility seems to be persistent to some extent, a change from low volatility to high will typically last for more than one observation.

This type of behaviour can be modeled by a stochastic process with the mean-reverting property, i.e. process with drift of the form $\mu(x,t) = \kappa(\theta - x)$. In this representation the parameter κ represents a measure of persistence and the parameter θ represents the long term mean.

An important property of volatility is that it is always positive. This is because negative volality makes no practical sense. The Bessel process has both of these properties which makes it a good candidate for modeling stochastic volatility. A Bessel process is

defined as a solution to stochastic differential equation

$$dx = \kappa(\theta - x)dx + \sigma\sqrt{x}dW. \tag{1.12}$$

1.5 Fong-Vasicek Model with Stochastic Volatility

The Fong-Vasicek model with stochastic volatility, first defined in [13], belongs to a class of two factor interest rate models, meaning the yield curve $R(\tau, r)$ is determined by it's origin, i.e. the short rate r. The stochastic process behind the short rate r depends on itself but also on a second factor, in this case volatility y. The second process is defined in line with the notes made in previous section, describing the character of volatility observed on empirical data. As a two-factor model, it offers more variety in terms of possible outcomes than one-factor models and as such is expected be a better fit for empirical data. The Fong-Vasicek model with stochastic volatility is specified by two stochastic differential equations that govern the behaviour of short rate r and volatility y, they are

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}dW_1 \\ dy &= \kappa_2(\theta_2 - y)dt + v\sqrt{y}dW_2. \end{aligned} \tag{1.13}$$

The model allows for constant correlation ρ between the increments dW_1 and dW_2, i.e. $\mathrm{E}(dW_1, dW_2) = \rho dt$. If the market prices of risk are $\lambda_1\sqrt{y}$ and $\lambda_2\sqrt{y}$ then the discount bond price function $P(\tau, r, y)$, where τ is time to maturity of the bond, solves a partial differential equation

$$\begin{aligned} &-\frac{\partial P}{\partial \tau} + (\kappa_1(\theta_1 - r) - \lambda_1 y)\frac{\partial P}{\partial r} + (\kappa_2(\theta_2 - y)\lambda_2 vy)\frac{\partial P}{\partial y} \\ &+\frac{y}{2}\frac{\partial^2 P}{\partial r^2} + \frac{v^2 y}{2}\frac{\partial^2 P}{\partial y^2} + \rho v y \frac{\partial^2 P}{\partial r \partial y} - rP = 0, \end{aligned} \tag{1.14}$$

with the initial condition $P(0, r, y) = 1$. This PDE is known to have a solution of the form

$$P(\tau, r, y) = A(\tau)e^{-B(\tau)r - C(\tau)y} \tag{1.15}$$

see e.g. [24] or [13]. Functions A, B and C are solutions to a system of ODE obtained from (1.14) by plugging in a solution of this form. The system of ODE reads as follows

$$\begin{aligned} A' &= -A(\kappa_1\theta_1 B + \kappa_2\theta_2 C), \\ B' &= -\kappa_1 B + 1, \\ C' &= -\lambda_1 B - \kappa_2 C - \lambda_2 v C - \frac{B^2}{2} - \frac{v^2 C^2}{2} - v\rho BC, \end{aligned} \tag{1.16}$$

with initial conditions $A(0) = 1$, $B(0) = 0$ and $C(0) = 0$, which correspond to the condition $P(0, r, y) = 1$. This system can be solved in steps. First B can be found analytically. This solution is then inserted into ODE for C, which is solved numerically, using the Runge-Kutta method. Both solutions are inserted into the integrated version of ODE for A to find the values of A. This scheme can be summarized by equations

$$B = \frac{1}{\kappa_1}(1 - e^{-\kappa_1\tau}), \tag{1.17}$$

$$C' = -\lambda_1 B - \kappa_2 C - \lambda_2 v C - \frac{B^2}{2} - \frac{v^2 C^2}{2} - v\rho BC, C(0) = 0, \tag{1.18}$$

$$A = \exp\left(-\theta_1\tau + \theta_1 B - \kappa_2\theta_2 \int_0^\tau C(s)ds\right). \tag{1.19}$$

Knowing how to numerically evaluate functions A, B and C we take the values of parameters used in [20], which can be found in Table 1.1, to compare our results. These parameters are estimated from empirical data in work [10]. The functions A, B, C for these values of parameters are plotted, with large maturities that allow us to observe their limit behaviour, in Figure 1.3.

κ_1	θ_1	κ_2	θ_2	v	ρ	λ_1	λ_2
0.109	0.0652	1.482	0.000264	0.01934	0	-11	-6

Table 1.1: Baseline parameter values, previously used in [20]

We also know that the function A, B, C have well defined limits for $\tau \to \infty$, given that a structural condition

$$\frac{1 + 2\lambda_1\kappa_1}{2\kappa_1} < 0 \tag{1.20}$$

is satisfied. For our values of parameters this in this case. We know that the domain of function A is the interval $(0, 1)$, function B has a limit equal to $1/\kappa_1 = 9.17$ in

this case, this limit can be observed directly from 1.17. In [20] it is established that if function C has a limit $C_L \geq 0$, then C_L is the single positive root of the quadratic equation

$$0 = \frac{v^2}{2}C_L^2 + (\kappa_2 + \lambda_2 v + \frac{v\rho}{\kappa_1})C_L + \frac{1 + 2\kappa_1}{2\kappa_1^2},$$

given that the structural condition (1.20) is satisfied. For our parameters this is the case and the positive root of this equation is approximately $C_L = 42.82$. Figure 1.3 confirms the derived limit behaviour for all three functions.

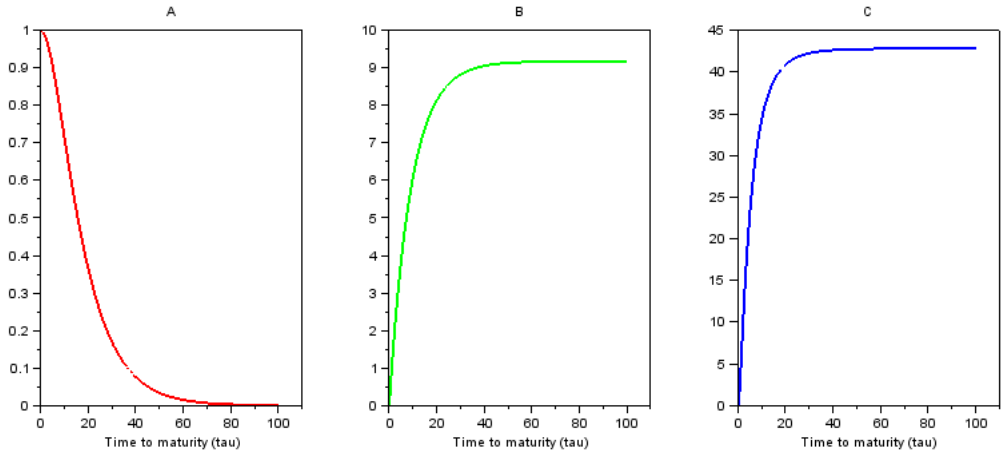


Figure 1.3: Functions A, B, C, with parameters taken from Table 1.1

2 Asymptotic Approximation of Bond Price in Fong-Vasicek Model

2.1 The fast scale of volatility

In empirical work, see e.g. [9] or [14], it has been observed that the volatility process y operates on a different time scale than the short term process r. Namely the scale of the volatility process is much faster. Fast scale volatility can be incorporated into the model with stochastic volatility. We present two different approaches of doing so, while assuming that the volatility process y is a Bessel process defined by equation (1.12).

The first approach, used in [25], is to rewrite the defining equation into

$$dy = \frac{\kappa(\theta - y)}{\epsilon}dt + \frac{\sigma\sqrt{y}}{\sqrt{\epsilon}}dW. \tag{2.1}$$

This is a result of considering the volatility process y in scale of $t\epsilon$ instead of the usual t. Here the parameter ϵ should be a positive number in the neighbourhood of zero.

The second approach, defined in [14] and used in [9] and [20], is to say that the fast scale implies a large value of κ in the equation (1.12) and define the unit of fast scale $\epsilon = 1/\kappa$. Making this substitution into equation (1.12) to obtain

$$dy = \frac{(\theta - y)}{\epsilon}dt + \frac{\sigma\sqrt{y}}{\sqrt{\epsilon}}dW. \tag{2.2}$$

The term defining the volatility of y changes by a factor depending on ϵ to keep the variance of the limit distribtion of y constant, for more details see [20].

Cotton et. al. in [9] utilize the fast scale of volatility to approximate two factor short rate models by a one factor model plus a correction term. The main result of their paper is a theorem proving that the discount bond pricing function P can be approximated by a function P^ϵ with an error in $O(\epsilon)$. The remarkable thing about this approximation is the fact that it does not depend on the value of volatility y. This is a very attractive feature as the value of y is not directly observable in practice. Not only that but in terms of fitting the function P^ϵ has less parameters to fit. The proposed idea is to utilize the correction term to provide additional space for better

fitting, while keeping the number of parameters to be estimated lower than in a two factor model. This method of asymptotic approximation for two factor Fong-Vasicek model was derived in [20].

2.2 Asymptotic approximation of the bond price in Fong-Vasicek model

In this section we present the results of [20] concerning the approximation of bond pricing function P by a function P^ϵ. This approximation is based on the idea of fast scaling volatility.

The model we work with is the Fong-Vasicek model with stochastic volatility described previously in section 1.5. We take the second approach to incorporating fast scaling volatility. In this approach the fast scale of volatility is translated into the parameters in the form of assuming a large value of κ_2, which inverted becomes the unit of fast scale, the parameter $\epsilon = 1/\kappa_2$. The original defining equations for Fong-Vasicek model (2.3), are adapted to keep the variance of the asymptotic distribution of y process constant and change into

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}dW_1 \\ dy &= \frac{1}{\epsilon}(\theta_2 - y)dt + \frac{v\sqrt{y}}{\sqrt{\epsilon}}dW_2, \end{aligned} \tag{2.3}$$

where W_1 and W_2 are Wiener processes and $\mathrm{E}(dW_1, dW_2) = \rho$.

We assume that the market prices of risk can be written as $\lambda_1\sqrt{y}$ and $\lambda_2\sqrt{y}$, with λ_1 and λ_2 constant. Then in this model the bond pricing function P^ϵ is a solution to the PDR

$$\begin{aligned} \frac{\partial P^\epsilon}{\partial t} + (\kappa_1(\theta_1 - r) - \lambda_1 y)\frac{\partial P^\epsilon}{\partial r} + \left(\frac{1}{\epsilon}(\theta_2 - y) - \frac{1}{\sqrt{\epsilon}}\lambda_2 vy\right)\frac{\partial P^\epsilon}{\partial y} \\ + \frac{y}{2}\frac{\partial^2 P^\epsilon}{\partial r^2} + \frac{1}{\epsilon}\frac{v^2 y}{2}\frac{\partial^2 P^\epsilon}{\partial y^2} + \frac{1}{\sqrt{\epsilon}}\rho v y \frac{\partial^2 P^\epsilon}{\partial r \partial y} - rP^\epsilon = 0, \end{aligned} \tag{2.4}$$

with the end condition $P^\epsilon(T, r, y) = 1$.

The pricing function P^ϵ can be rewritten as an infinite sum, an asymptotic series based on powers of $\sqrt{\epsilon}$,

$$P^\epsilon = P_0 + \sqrt{\epsilon}P_1 + \epsilon P_2 + \epsilon\sqrt{\epsilon}P_3 + ..., \tag{2.5}$$

where P_i are functions of t, r and y and they are subject to conditions $P_0(T, r, y) = 1$ and $P_i(T, r, y) = 0 \quad \forall i \geq 1$. In this series a cut-off point is chosen. The terms left before the cut-off point, are then the asymptotic approximation of P^ϵ. Seleceniova in [20] shows this by implementing a clever use of operator functions, also used in [9], which make the calculations possible. Subsequently the individual terms of the asymptotic series are calculated and conditions for them formulated. Ultimately the result is that the price of the discount bond can be approximated as

$$P^\epsilon(\tau, r, y) \approx P_0(\tau, r) + \sqrt{\epsilon} P_1(\tau, r), \tag{2.6}$$

where

$$\begin{aligned} P_0(\tau, r) &= A(\tau) e^{-B(\tau) r}, \\ P_1(\tau, r) &= D(\tau) A(\tau) e^{-B(\tau) r}, \end{aligned}$$

and the functions B, A and D are defined as

$$\begin{aligned} B(\tau) &= \frac{1}{\kappa_1}(1 - e^{-\kappa_1 \tau}), \\ A(\tau) &= \exp\left[(B(\tau) - \tau)(\theta_1 - \frac{\lambda_1 \theta_2}{\kappa_1} - \frac{\theta_2}{2\kappa_1^2}\right) - \frac{\theta_2}{4\kappa_1} B(\tau)^2\Big], \\ D(\tau) &= \frac{V_1}{\kappa_1}(-B(\tau) + \tau) - \frac{V_2}{\kappa_1^2}(-B(\tau) - \frac{\kappa_1}{2} B(\tau)^2 + \tau) \\ &\quad + \frac{V_3}{\kappa_1^3}(-B(\tau) - \frac{\kappa_1}{2} B(\tau)^2 - \frac{\kappa_1^2}{3} B(\tau)^3 + \tau). \end{aligned} \tag{2.7}$$

The group parameters V_1, V_2 and V_3 arise during the process of deriving the approximation. They can be expressed, in terms of model parameters, as

$$\begin{aligned} V_1 &= -\lambda_1 \lambda_2 v \theta_2, \\ V_2 &= \frac{1}{2} \lambda_2 v \theta_2 + \lambda_1 \rho v \theta_2, \\ V_3 &= -\frac{1}{2} \rho v \theta_2. \end{aligned}$$

The main difference between this approach and the original Fong-Vasicek model with stochastic volatility is in the fact that the functions P_0 and P_1 do not depend on y. This is a desirable property, because in practice the current value of volatility is unknown and cannot be directly observed.

Another attractive feature of this is approximation is the decreased number of parameter that need to be estimated to fit this model. In the original Fong-Vasicek model we have parameters v and ρ and also the parameters λ_1, λ_2 tied with the market value of risk, which would need to be estimated. In the approximation these are replaced by group parameters V_1, V_2, V_3. These are closely linked to the parameters of the original model, but are fewer in number. Moreover [9] show that they can be estimated from yield curves. This gives us an advantage when estimating all parameters needed for pricing of bonds.

To illustrate the feasibility of this approximation of the exact bond price function P, defined by 1.15 we plot the functions P and P^ϵ together for values of $r = 0.04$, $r = 0.08$ and five values of volatility from range $y = [0.00001, 0.00011]$. The results can be seen in Figure 2.1 and indicate that P^ϵ is a good aproximation of P as the plots are close in both cases. To elucidate the difference we give their relative difference defined as $(P(\tau, r, y) - P^\epsilon(\tau, r))/(P(\tau, r, y)$ in Figure 2.2. We can see that this relative difference increases with time to maturity τ, at least in the time frame considered, but they remain well within accepted range.

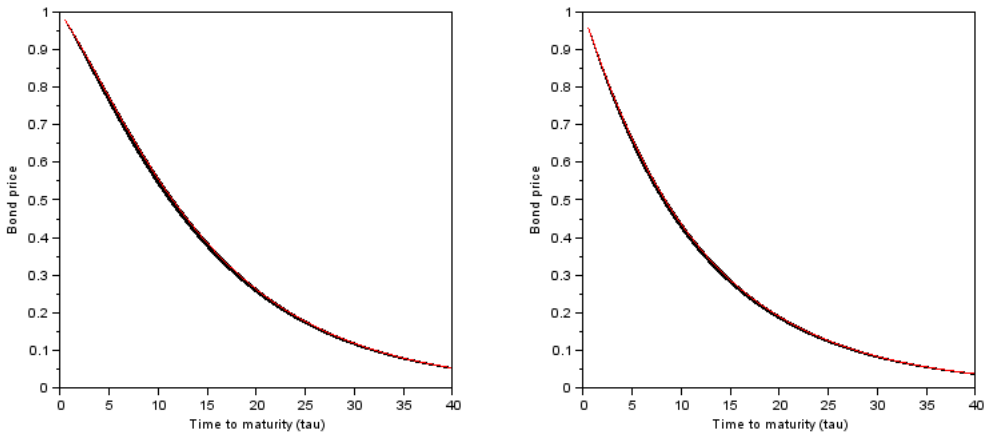


Figure 2.1: Comparison of bond pricing function P and it's approximation P^ϵ for $r = 0.04$ on the left and $r = 0.08$ on the right. Volatility is taken from range $[0.0001, 0.0011]$ in both cases.

In our work we will not be interested directly in the price of the bonds but rather

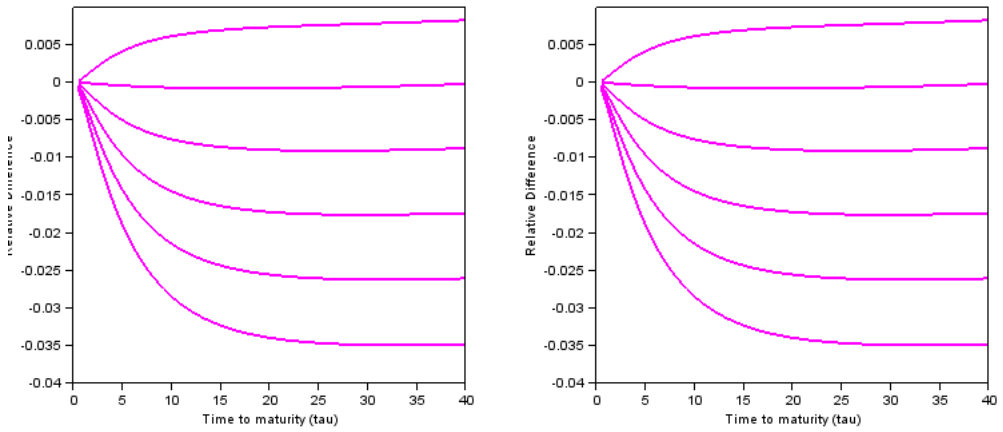


Figure 2.2: Relative difference between the bond pricing function P and it's approximation P^ϵ, defined as $(P(\tau, r, y) - P^\epsilon(\tau, r))/(P(\tau, r, y)$, for $r = 0.04$ on the left and $r = 0.08$ on the right. Volatility is taken from range $[0.0001, 0.0011]$ in both cases.

we will be working with yield curves, defined by

$$R(\tau, r, y) = -\frac{ln(P(\tau, r, y))}{\tau}. \tag{2.8}$$

We will approximate the yield curves by

$$R^\epsilon(\tau, r) = -\frac{ln(P^\epsilon(\tau, r))}{\tau}. \tag{2.9}$$

Figure 2.3 shows the exact yield curves R for two different values of short rate r and a range of different values of volatility y. These are compared to the approximation R^ϵ, in red. The approximation is the same for all the different values of volatility y, since it does not depend on the its current value.

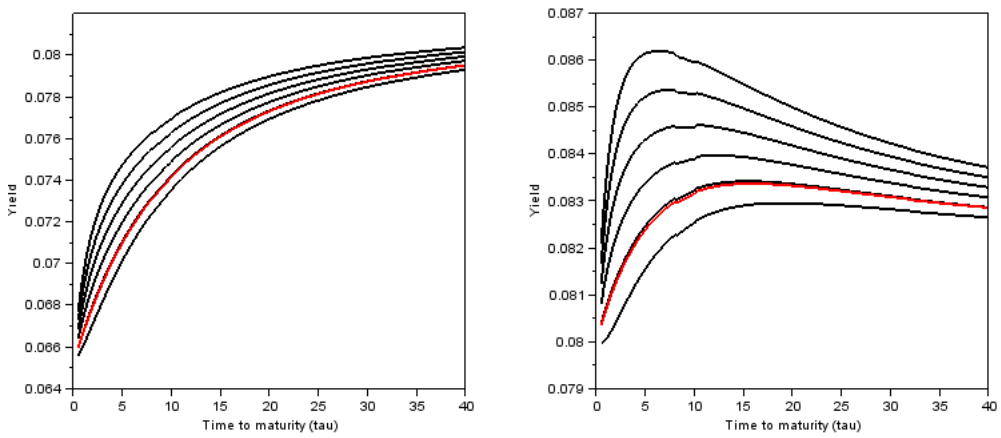


Figure 2.3: Yield curves R, in black, and their approximation R^ϵ, in red, for $r = 0.0652$ on the left and $r = 0.08$ on the right. Volatility y is taken from range $[0.0001, 0.0011]$ in both cases.

3 Generating yield curve data following Fong-Vasicek model

3.1 Generating short rate and volatility trajectories

In Fong-Vasicek model with stochastic volatility the short rate r and volatility y follow a stochastic process specified by a set of differential equations

$$\begin{aligned} dr &= \kappa_1(\theta_1 - r)dt + \sqrt{y}dW_1 \\ dy &= \kappa_2(\theta_2 - y)dt + v\sqrt{y}dW_2, \end{aligned} \tag{3.1}$$

where we allow for correlation ρ between the random terms dW_1 and dW_2. To generate trajectories we discretize these continuous differential equations in the following way. Considering the definition of dr it is the change in short rate r that happens with an infinitesimal change in time t. We choose to approximate this change by

$$dr \approx r(t + \Delta) - r(t), \tag{3.2}$$

where Δ should be small and use this definition as the basis of the discretization. We pick a time step $\Delta = 0.01$, which we deem to be sufficiently small. With value of Δ defined we have the ability rewrite the equations (3.1) into

$$\begin{aligned} r(t + \Delta) &= r(t) + \kappa_1(\theta_1 - r)\Delta + \sqrt{y}W_1, \\ y(t + \Delta) &= y(t) + \kappa_2(\theta_2 - y)\Delta + v\sqrt{y}W_2, \end{aligned} \tag{3.3}$$

where the terms W_1 and W_2 have a distribution

$$(W_1, W_2)^T \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Delta \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right). \tag{3.4}$$

The data generating process starts by picking suitable starting values. In this work the starting values of process r and y are taken to be their long term averages, i.e. $r(0) = \theta_1$ and $y(0) = \theta_2$. Then we generate a fixed number of observations from the distribution of $W = (W_1 W_2)^T$. We do this by generating two independ samples from standard normal distribution and multiplying the by the square root of the covariance

matrix to obtain a randomly generated vector with distribution (3.4). We then generate observations of processes r and y iteratively utilizing equations (3.3). We generate a sample of predetermined length N, which varies by application. Because it would not random enough for our purposes to have all trajectories starting from the same starting points we set a lower limit l, usualy $l = 100$, and consider only the sample strating from $(r(l), y(l)$.

3.2 Parameter sets used to generate data in this work

The question of what parameters to use for generating data from Fong-Vasicek model is an important one. Authors in [23] show that certain intuitive conditions imposed on the bond price P are satisfied if the following structural condition holds true

$$\lambda_1 \leq -\frac{1}{2\kappa_1}. \tag{3.5}$$

This is a condition that the parameter sets we use all satisfy.

The work [11] deals with the estimation of parameters of the Fong-Vasicek model from empirical data. In [20] the parameters estimated in [11] are revisited and refined. The first parameter set we use are the parameters described in [20] as the realisticly best fit. These parameters were already mentioned and can be found in Table 1.1. In our work we consider more than a single set of parameters. In the second parameter set we consider, parameter set 2, we introduce positive correlation between the increments dW_1 and dW_2 in 3.1. The third and fourth parameter sets are the same as the first with change in parameter κ_2. We know that κ_2 is a parameter that is very closely related to the approximation described in section 2.2. The inverted value of this paramter is the unit of the fast scaling volatility $\epsilon = 1/\kappa_2$. For reasons explained in the said section the approximation should be more precise for large values of κ_2. This is the reason why we take κ_2 ten times larger in the second pair of parameter sets. The last parameter set is used to generate data from one factor Vasicek model. Notice that for these values of parameters the volatility process y has zero increments dy and is thus always equal to its starting value θ_2. Table 3.1 is a summary of all parameter sets used to generate the data from Fong-Vasicek model throughout this work.

	κ_1	θ_1	κ_2	θ_2	v	ρ	λ_1	λ_2
Parameter set 1	0.109	0.0652	1.482	0.000264	0.01934	0	-11	-6
Parameter set 2	0.109	0.0652	1.482	0.000264	0.01934	0.7	-11	-6
Parameter set 3	0.109	0.0652	14.82	0.000264	0.01934	0	-11	-6
Parameter set 4	0.109	0.0652	14.82	0.000264	0.01934	0.7	-11	-6
Parameter set 5	0.109	0.0652	0	0.000264	0	0	0	0

Table 3.1: Parameter sets used in this work to generate data from Fong-Vasicek model, the first parameter set is taken from [20]. The others are its derivations.

A note should be made here about measures. The data generating process uses the parameters under risk neutral measure. The estimation procedure defined in the next section estimates the parameters under real measure. The difference is that under the real measure the parameters λ_1 and λ_2 are implicit and are part of the other parameters, which change values under different measures. For more information see e.g. [18].

To illustrate the impact of the changes in parameters we include Figures 3.1, generated using parameter set 1, and 3.2, generated using parameter set 4. Both figures exhibit the mean-reversion property in both factors. This is the property of the underlying model and should come as no surprise. Notice however that the trajectories of volatility are dramatically different. These figures illustrate very clearly the meaning of fast scaling volatility. The unit of fast scaling volatility, defined as

$$\epsilon = \frac{1}{\kappa_2},$$

is 10 times lower in parameter set 4. This results in much shorter periods away from the mean.

3.3 Generating yield curve data

In section 3.1 we derived a method of generating trajectories of short rate r and volatility y that follow the Fong-Vasicek model. These processes define the yield curves $R(t, r, y)$. For the purposes of this work we need to generate these yield curves. We will use the generated data as observations onto which we will fit the Vasicek model

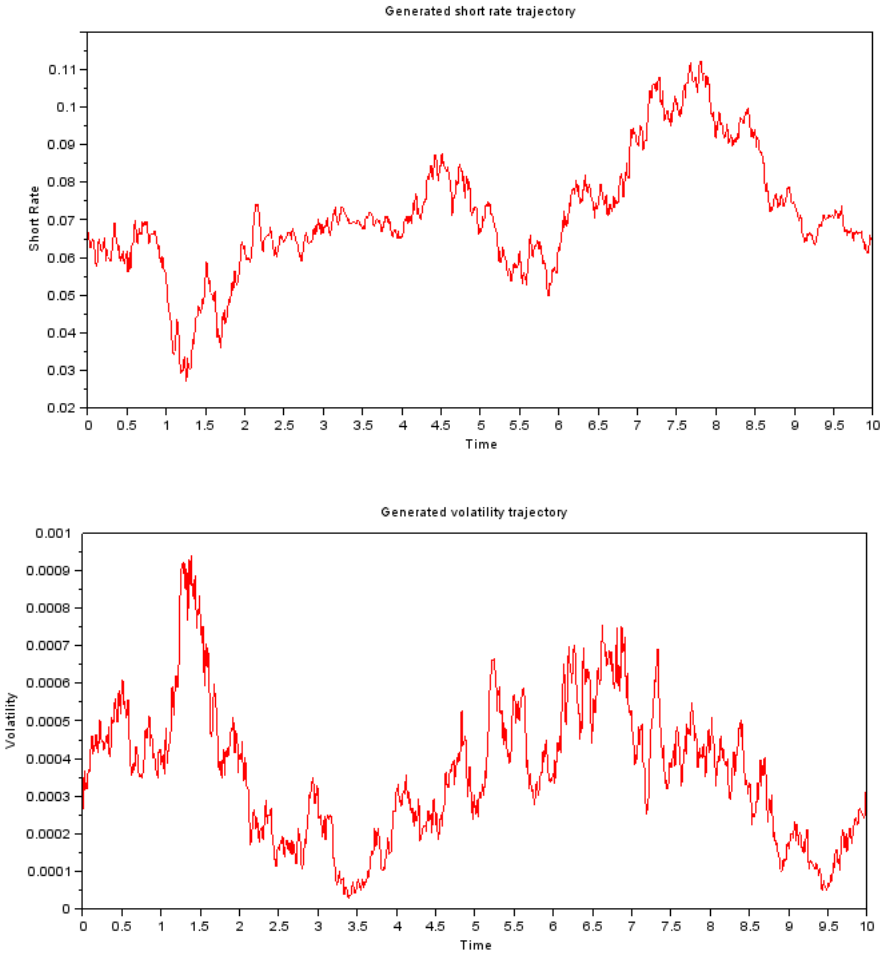


Figure 3.1: A generated trajectory for the two factors of Fong-Vasicek model, the short rate r and volatility y. This figure was generated using parameter set 1.

and the asymptotic approximation of Fong-Vasicek model. We know that in the Fong-Vasicek model, the price of a bond P is defined by the equation

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

where the functions A, B and C solve the system of ODE (1.16). This system depends on the value of parameters of the Fong-Vasicek model. We solve this system numericaly, employing the Runge-Kutta method, for a given set of parameters. The yield curves R are defined, in terms of prices P, as

$$R(\tau, r, y) = -\frac{\ln P(\tau, r, y)}{\tau}.$$

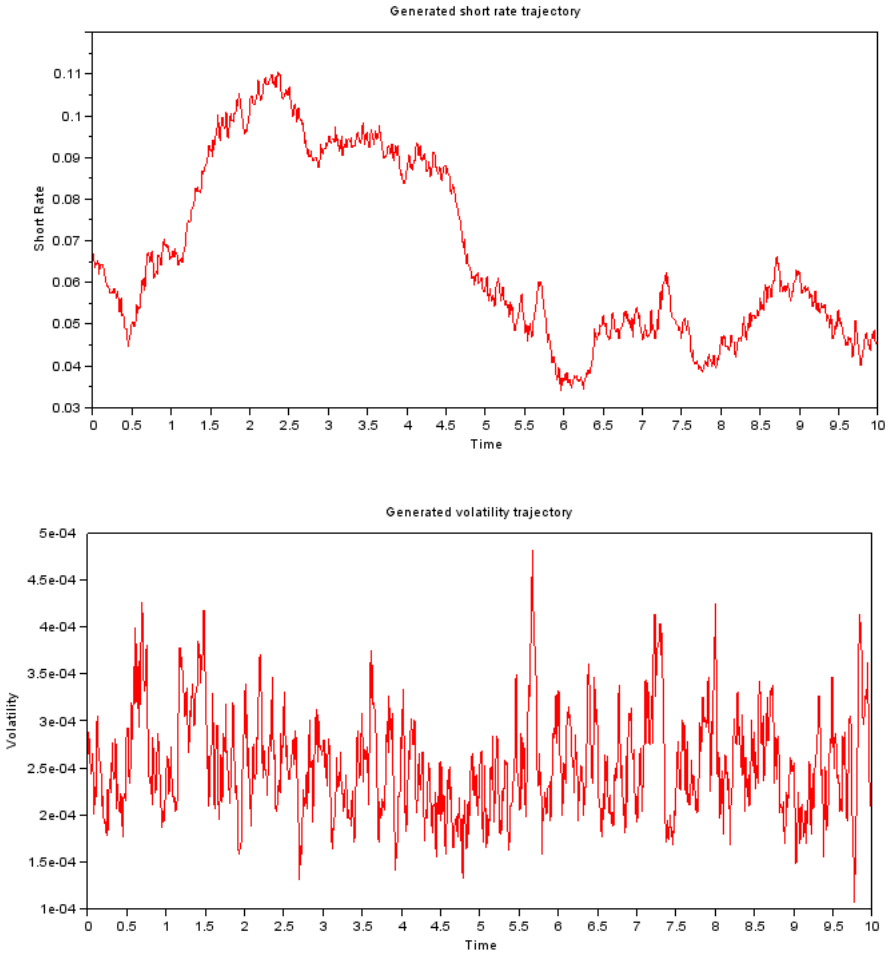


Figure 3.2: A generated trajectory for the two factors of Fong-Vasicek model, the short rate r and volatility y. This figure was generated using parameter set 4.

We generate the yield curves for consecutive days by using the numericaly solved functions A, B, C and the observations of r and y generated as specified in section 3.1. The generated yield curves are then stored in a variable if they are used immediately or in a text file, with a reasonably defined structure, for later use in numerical experiments. Below are the first few lines of a text file generated by this procedure. To give an idea about the shape of the yield curves using this data generating process we include Figure 3.3, which contains a few yield curves generated using parameter set 4.

```
Following is a table of generated daily yield curves with maturities in years defined on the first row.
    0.25,     0.50,        1,        2,        3,        4,        5,        6,        7,        8,        9,       10,       20,       30
 0.06035,  0.06071,  0.06145,  0.06290,  0.06425,  0.06548,  0.06659,  0.06759,  0.06850,  0.06932,  0.07002,  0.07069,  0.07502,  0.07708
 0.06050,  0.06087,  0.06161,  0.06306,  0.06441,  0.06563,  0.06673,  0.06773,  0.06863,  0.06944,  0.07013,  0.07080,  0.07510,  0.07713
 0.06119,  0.06156,  0.06230,  0.06372,  0.06504,  0.06623,  0.06731,  0.06828,  0.06915,  0.06994,  0.07061,  0.07125,  0.07539,  0.07735
 0.06130,  0.06167,  0.06241,  0.06384,  0.06516,  0.06635,  0.06742,  0.06838,  0.06925,  0.07003,  0.07070,  0.07134,  0.07545,  0.07739
```

```
0.06219,  0.06256,  0.06329,  0.06468,  0.06596,  0.06711,  0.06814,  0.06907,  0.06990,  0.07066,  0.07129,  0.07191,  0.07583,  0.07766
0.06169,  0.06206,  0.06280,  0.06421,  0.06551,  0.06668,  0.06774,  0.06868,  0.06954,  0.07030,  0.07096,  0.07159,  0.07562,  0.07751
0.06138,  0.06174,  0.06247,  0.06389,  0.06520,  0.06638,  0.06745,  0.06841,  0.06928,  0.07006,  0.07072,  0.07137,  0.07547,  0.07740
0.06204,  0.06239,  0.06311,  0.06449,  0.06577,  0.06693,  0.06797,  0.06890,  0.06975,  0.07051,  0.07115,  0.07178,  0.07574,  0.07760
0.06147,  0.06183,  0.06255,  0.06395,  0.06525,  0.06644,  0.06750,  0.06846,  0.06932,  0.07010,  0.07077,  0.07141,  0.07550,  0.07742
```

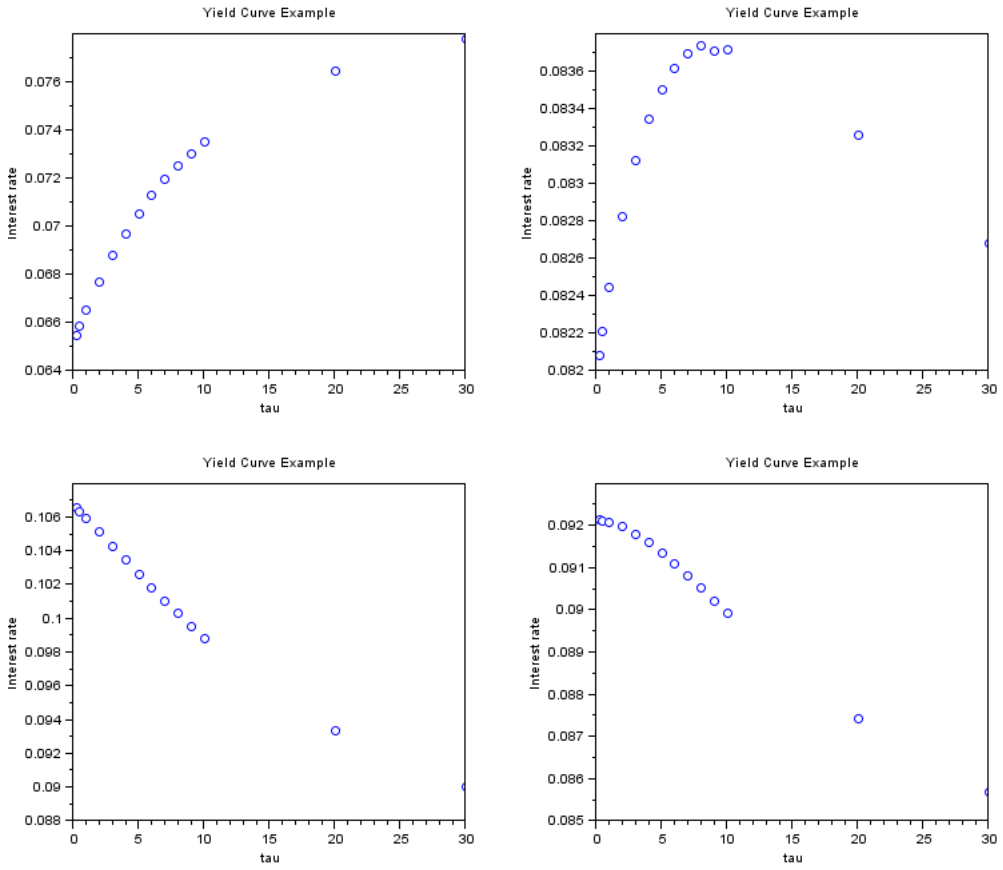


Figure 3.3: A few examples of the different shapes of yield curves generated by the data generating process defined in this section. Parameter set 4 from Table 3.1 was used to generate these data.

4 Estimating short rate models

The main goal of this work is to provide means of fitting the models defined in previous chapters onto data. In this section we describe a general framework we use to fit all models used in this work. The subsequent sections deal with issues of estimating the individual models.

4.1 General approach

In general there are two different approaches that can be used to fit a short rate model. The first option is fitting the stochastic process of the short rate. The data with a small enough maturity, e.g. 3 months, can be used as observations from the random process governing the short rate. Using for example the method of maximum likelihood one can find the ML estimators of the model parameters and fit the model onto data. In work [6], the authors employ the generalized method of moments applied on observations of the short rate to determine the properties of volatility. In this work we choose not to take this path. We could take the data with a small enough maturity as observations of the short rate process, however we have would have employ some means of construction observations of the volatility process y from available data.

The main focus of this work is to estimate the asymptotic approximation of Fong-Vasicek model which does not depend directly on the values of volatility process y. This brings us to the second approach, the one we will adopt. The idea behind it is that instead of fitting observations of the short rate r we fit the observed yield curves. This idea is used in the work [21]. The short rate models we use have a relatively simply defined discount bond price function P_M, with subscript M for modeled. This being the case it is very straightforward to calculate, given some values of model parameters and the short rate, what the yield curves R_M would be by a similar logic as was used in section 3.3. The fitting procedure we use to fit the model is minimizing the square difference between these calculated, or theoretical, yield curves

$$R_M = -\frac{\ln P_M(\tau_j, r_i)}{\tau_j},$$

and the observed data R. More explicitly, we define a cost function

$$F = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} w(i,j) \left[\frac{\ln P_M(\tau_j, r_i)}{\tau_j} + R_{i,j} \right]^2, \tag{4.1}$$

where we also allow for a function of weights $w(i,j)$. The parameters that minimize function F are our estimators of the model parameters governing the underlying process.

Numerical optimization of the function F is done in steps. The parameters are divided into two groups. The inner group, optimized in inner iteration, can be calculated based on the fixed values of parameters in the outer group. Their optimal values are calculated as solutions to first order conditions, FOC, minimizing the cost function F, considering the outer parameters as fixed constants. The values of the outer parameters are optimized using Newton's method, see e.g. [26]. The standard Newton's method starts with a parameter vector θ_n in each step and defines the next parameter vector as

$$\theta_{n+1} = \theta_n - H(F)^{-1}(\nabla F),$$

where $H(F)$ is the Hessian matrix of function F and ∇F is its gradient, both evaluated at θ_n. We approximate the Hessian matrix and the gradient by using the following logic. We approximate the individual elements of the gradient vector using

$$\begin{aligned} \frac{\partial F}{\partial x_i}(x) &= \lim_{h \to 0} \frac{F(x + e_i h) - F(x)}{h}, \\ &\approx \frac{F(x + e_i h) - F(x)}{h} \quad \text{evaluated at } h \text{ small enough}, \end{aligned}$$

where e_i denotes a vector of zeros with a single one in i-th position and what small enough means for the value of h is determined case by case. We use a very similar

notion to approximate the elements of $H(F)$ by taking

$$\begin{aligned}
\frac{\partial}{\partial x_j}\frac{\partial F}{\partial x_i}(x) &= \frac{\partial}{\partial x_j}\lim_{h\to 0}\frac{F(x+e_ih)-F(x)}{h},\\
&= \lim_{h\to 0}\frac{\partial}{\partial x_j}\frac{F(x+e_ih)-F(x)}{h},\\
&= \lim_{h\to 0}\frac{\frac{F(x+e_jh+e_ih)-F(x+e_jh)}{h} - \frac{F(x+e_ih)-F(x)}{h}}{h},\\
&= \lim_{h\to 0}\frac{F(x+e_jh+e_ih)-F(x)}{h^2} - \frac{1}{h}\left(\frac{\partial F}{\partial x_i}(x)+\frac{\partial F}{\partial x_j}(x)\right),\\
&\approx \frac{F(x+e_jh+e_ih)-F(x)}{h^2} - \frac{1}{h}\left(\frac{\partial F}{\partial x_i}(x)+\frac{\partial F}{\partial x_j}(x)\right),
\end{aligned}$$

evaluated at h small enough,

as before the value of h is determined case by case, but is taken the same for both approximations. To make the method computationaly efficient we reuse the elements of gradient already approximated in the calculation of the $H(F)$.

4.2 One Factor Vasicek Model

In this section we derive the estimation of a one factor Vasicek model. We will compare the fit of this model with the results from the asymptotic approximation of Fong-Vasicek model, defined in the next section. It can also serve to ilustrate the ideas of parameter estimation on a simpler model, as less space is taken up by technicalities.

As outlined in the beginning of the section we estimate the parameters of the process by minimizing the square distance between the observed and theoretical yield curves. Our data are yield curves from n days, each having m observations coresponding to different maturities. We denote the data $R_{i,j}$ for an interest rate from day i corresponding to j-th maturity. From (1.9) we know how to calculate the theoretical price of the discount bond $P(\tau, r)$, which defines the yield curve as

$$R(\tau, r) = -(\ln P(\tau, r))/\tau.$$

Plugging what we know about $P(\tau, r)$, from section 1.3, into the equation we obtain

$$\begin{aligned} R(\tau, r) &= -\frac{1}{\tau} \ln \left(A(\tau) e^{-B(\tau) r} \right) \\ &= -\frac{1}{\tau} \left[(\ln A(\tau)) - B(\tau) r \right] \\ &= -\frac{1}{\tau} \left[(-\theta + \frac{\sigma^2}{2\kappa^2})(-\frac{1}{\kappa}(1 - e^{-\kappa\tau}) + \tau) - \frac{\sigma^2}{4\kappa^3}(1 - e^{-\kappa\tau})^2 - r\frac{(1 - e^{-\kappa\tau})}{\kappa} \right]. \end{aligned}$$

While this expression is far from simple it can be rearanged, as shown in [7], into a simpler form if we fix the value of parameter κ. We can then obtain the following expression

$$R(\tau, r) = -(c_0(\tau, r) + c_1(\tau, r)\theta + c_2(\tau, r)\sigma^2), \tag{4.2}$$

where the coefficient functions c_i are

$$\begin{aligned} c_0(\tau, r) &= -r\frac{(1 - e^{-\kappa\tau})}{\kappa\tau} \\ c_1(\tau, r) &= \frac{1}{\tau}(\tau - \frac{1}{\kappa}(1 - e^{-\kappa\tau})) \\ c_2(\tau, r) &= \frac{1}{2\kappa^2\tau} \left(\tau - \frac{1}{\kappa}(1 - e^{-\kappa\tau}) - \frac{1}{2\kappa}(1 - e^{-\kappa\tau})^2 \right). \end{aligned}$$

It is important to notice that all these coefficient functions also depend on the fixed value of κ.

Having a managable expression for the theoretical yield curves $R(\tau, r)$ we can define cost function F, measuring the square difference between the observed and theoretical yield curves,

$$F = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} w(i, j) \left(R_{i,j} + c_0(\tau_j, r_i) + c_1(\tau_j, r_i)\theta + c_2(\tau_j, r_i)\sigma^2 \right)^2. \tag{4.3}$$

In our work we use $w(i, j) = \tau_j^2$, where τ_j is the j-th maturity observed in the data, these values are taken the same as in [7]. Note that function F depends on values of three parameters, namely θ, σ^2 and κ. In the inner step we consider κ to be a fixed value. Observing the first order conditions with respect to θ and σ^2 we find that we can eliminate these from optimization as their optimal values can be calculated from

FOC. Consider the FOC for the inner problem:

$$\frac{\partial F}{\partial \theta} = 0,$$
$$\frac{\partial F}{\partial \sigma^2} = 0.$$

If we substitute for F its definition from (4.3), these conditions become

$$\hat{\theta}\left(\sum_{i,j} c_1(\tau_j, r_i)^2 w_{i,j}\right) + \hat{\sigma^2}\left(\sum_{i,j} c_1(\tau_j, r_i)c_2(\tau_j, r_i)w_{i,j}\right) = -\left(\sum_{i,j} c_1(\tau_j, r_i)w_{i,j}(c_0(\tau_j, r_i) + R_{i,j})\right),$$
$$\hat{\theta}\left(\sum_{i,j} c_1(\tau_j, r_i)c_2(\tau_j, r_i)w_{i,j}\right) + \hat{\sigma^2}\left(\sum_{i,j} c_2(\tau_j, r_i)^2 w_{i,j}\right) = -\left(\sum_{i,j} c_2(\tau_j, r_i)w_{i,j}(c_0(\tau_j, r_i) + R_{i,j})\right).$$

Notice that the terms in from of the parameter estimators and the factors on the right hand side of these equations do not depend on them. To simplify the looks of this system of linear equations we rewrite them as

$$\hat{\theta}a_1 + \hat{\sigma^2}a_2 = a_3$$
$$\hat{\theta}a_4 + \hat{\sigma^2}a_5 = a_6.$$

Thus we have a system of two equations from which we can calculate values $\hat{\theta}$ and $\hat{\sigma^2}$ as

$$\hat{\theta} = \frac{a_2a_6 - a_3a_5}{a_2a_4 - a_1a_5},$$
$$\hat{\sigma^2} = \frac{a_1a_6 - a_3a_4}{a_1a_5 - a_2a_4}.$$

These equations give us a way to calculate the optimal values of parameters given the fixed value of κ.

The outer optimization problem, is the one of finding the optimal value κ. This is a one dimensional optimization problem

$$\min_{\kappa} F = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} w(i,j) \left(R_{i,j} + c_0(\tau_j, r_i) + c_1(\tau_j, r_i)\hat{\theta} + c_2(\tau_j, r_i)\hat{\sigma^2} \right)^2 . \tag{4.4}$$

We solve this problem using the modification of Newton's method outlined in the beginning of this section. We use the method until a desired prescision in the value of parameter κ is obtained. In our work the desired precision was almost exclusively 10^{-4}.

We confirm the validity of this procedure by fitting data generated using the parameter set 5, as defined in Table 3.1. In the Table 4.1 we present the comparison of the parameters, translated into the notation of Vasicek model, used to generate the data and the estimates of these parameters. We can see that the fit is almost perfect. We attribute the slight discrepancy simply to a limited number of observations. In this case 250 observations were generated as a sample and fitted. To illustrate the goodness of this we also include Figure 4.1 in which we present the fit of three particular yield curves, each representing one day out of $n = 250$, with dots denoting the generated data and blue line denoting the fitted yield curve.

	κ	θ	σ^2
used to generate	0.109	0.06520	0.0002460
estimated	0.10678	0.06488	0.0002476

Table 4.1: Vasicek Model Parameter Fitting

4.3 Asymptotic approximation of the Fong-Vasicek model

This section is focused on fitting the parameters of the asymptotic approximation of Fong-Vasicek model bond price P described in section 2.2. We again aim to minimize the distance between the observed yield curves and the theoretical ones in terms of the cost function

$$F = \frac{1}{mn} \sum_{i=1}^{n} \sum_{j=1}^{m} w(i,j) \left[\frac{\ln P(\tau_j, r_i)}{\tau_j} + R_{i,j} \right]^2 .$$

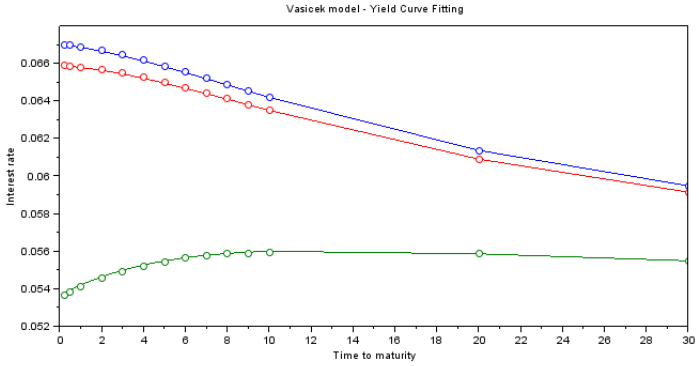


Figure 4.1: Fit of three particular yield curves, i.e. three days out of n, with dots denoting data and lines denoting the fitted yield curves.

The price of the discount bond under the assumptions of the approximation is defined in (2.6). Here we derive how this can be linearized as a function of inner group of parameters, while taking the outer parameters fixed.

Starting from the equation (2.6) consider the following

$$\begin{aligned}\ln P(\tau, r) &= \ln\left(P_0(\tau, r) + \sqrt{\epsilon} P_1(\tau, r)\right) \\ &= \ln\left[A(\tau)e^{-B(\tau)r} + \sqrt{\epsilon} D(\tau) A(\tau) e^{-B(\tau)r}\right] \\ &= \ln\left[A(\tau)e^{-B(\tau)r}\left(1 + \sqrt{\epsilon} D(\tau)\right)\right] \\ &= \ln(A(\tau)) - B(\tau)r + \ln\left(1 + \sqrt{\epsilon} D(\tau)\right).\end{aligned}$$

As a next step we approximate

$$\ln\left(1 + \sqrt{\epsilon} D(\tau)\right) \approx \sqrt{\epsilon} D(\tau),$$

which is the first term in the Taylor series of the left hand side. We consider this approximation good enough as all other terms of the Taylor series contain ϵ in power of at least 1 and the accuracy of the original approximation is $O(\epsilon)$, meaning this approximation does not make it much less acurate. Thus we arrive at

$$\ln P(\tau, r) = \ln(A(\tau)) - B(\tau)r + \sqrt{\epsilon} D(\tau). \tag{4.5}$$

Functions B, A and D are defined in (2.7) as

$$\begin{aligned}
B(\tau) &= \frac{1}{\kappa_1}(1 - e^{-\kappa_1 \tau}), \\
A(\tau) &= \exp\left[(B(\tau) - \tau)(\,\theta_1 - \frac{\lambda_1 \theta_2}{\kappa_1} - \frac{\theta_2}{2\kappa_1^2}\,) - \frac{\theta_2}{4\kappa_1} B(\tau)^2\right], \\
D(\tau) &= \frac{V_1}{\kappa_1}(-B(\tau) + \tau) - \frac{V_2}{\kappa_1^2}(-B(\tau) - \frac{\kappa_1}{2} B(\tau)^2 + \tau) \\
&\quad + \frac{V_3}{\kappa_1^3}(-B(\tau) - \frac{\kappa_1}{2} B(\tau)^2 - \frac{\kappa_1^2}{3} B(\tau)^3 + \tau).
\end{aligned}$$

If we take parameters κ_1, θ_2 and ϵ as fixed then $(\ln P(\tau, r))/\tau$ can be expressed as a linear function of the remaining parameters, i.e.

$$\frac{\ln P(\tau, r)}{\tau} = c_0(\tau, r) + c_1(\tau) V_1 + c_2(\tau) V_2 + c_3(\tau) V_3 + c_4(\tau)\theta_1 + c_5(\tau)\lambda_1. \tag{4.6}$$

Using the definitions of functions B, A, D and the equation (4.5) we can determine the coefficient functions to be

$$\begin{aligned}
c_0(\tau, r) &= \frac{1}{\tau}\left(-B(\tau) r - \frac{\theta_2}{2\kappa_1^2}(B(\tau) - \tau) - \frac{\theta_2}{4\kappa_1} B(\tau)^2\right), \\
c_1(\tau) &= \frac{\sqrt{\epsilon}}{\tau \kappa_1}(-B(\tau) + \tau), \\
c_2(\tau) &= \frac{\sqrt{\epsilon}}{\tau \kappa_1^2}\left(-B(\tau) - \frac{\kappa_1}{2} B(\tau)^2 + \tau\right), \\
c_3(\tau) &= \frac{\sqrt{\epsilon}}{\tau \kappa_1^3}\left(-B(\tau) - \frac{\kappa_1}{2} B(\tau)^2 - \frac{\kappa_1^2}{3} B(\tau)^3 + \tau\right), \\
c_4(\tau) &= \frac{1}{\tau}(B(\tau) - \tau), \\
c_5(\tau) &= -\frac{1}{\tau}\frac{\theta_2}{\kappa_1}(B(\tau) - \tau).
\end{aligned}$$

Using this linearization of $(\ln P(\tau, r))/\tau$ we could define the cost function F, measuring the difference between theoretical and observed yield curves. Notice however that the coefficient functions c_1, c_4 and c_5 are all constant linear transformations of the term $\frac{1}{\tau}(B(\tau) - \tau)$. Constant in this case meaning with respect to fixed parameters $\sqrt{\epsilon}, \theta_2$ and κ_1. As a direct result of this linear dependance the system of FOC is singular implying that the problem is overidentified, i.e. we have too many equations given the information contained in our data. We work around this problem by grouping all the colinear parameters into only one, i.e. we define

$$\bar{V_1} = -\frac{1}{\kappa_1} + \frac{\theta_1}{\sqrt{\epsilon}} - \frac{\theta_2}{\sqrt{\epsilon}\kappa_1}\lambda_1. \tag{4.7}$$

Using the new parameter $\bar{V_1}$, the linearization of $(\ln P(\tau, r))/\tau$ changes accordingly into

$$\frac{\ln P(\tau, r)}{\tau} = \gamma_0(\tau, r) + \gamma_1(\tau)\bar{V_1} + \gamma_2(\tau)V_2 + \gamma_3(\tau)V_3, \tag{4.8}$$

where

$$\begin{aligned}
\gamma_0(\tau, r) &= \frac{1}{\tau}\left(-B(\tau)r - \frac{\theta_2}{2\kappa_1^2}(B(\tau) - \tau) - \frac{\theta_2}{4\kappa_1}B(\tau)^2\right), \\
\gamma_1(\tau) &= \frac{\sqrt{\epsilon}}{\tau\kappa_1}(B(\tau) - \tau), \\
\gamma_2(\tau) &= \frac{\sqrt{\epsilon}}{\tau\kappa_1^2}\left(-B(\tau) - \frac{\kappa_1}{2}B(\tau)^2 + \tau\right), \\
\gamma_3(\tau) &= \frac{\sqrt{\epsilon}}{\tau\kappa_1^3}\left(-B(\tau) - \frac{\kappa_1}{2}B(\tau)^2 - \frac{\kappa_1^2}{3}B(\tau)^3 + \tau\right).
\end{aligned}$$

This allows to define the cost function F for this estimation problem as

$$F = \frac{1}{mn}\sum_{i=1}^{n}\sum_{j=1}^{m} w(i,j)\left(R_{i,j} + \gamma_0(\tau, r) + \gamma_1(\tau)\bar{V_1} + \gamma_2(\tau)V_2 + \gamma_3(\tau)V_3\right)^2, \tag{4.9}$$

where $w(i,j)$ are weight functions. We use the same weights as in the case of Vasicek model $w(i,j) = \tau_j^2$. The definition of the new group parameter $\bar{V_1}$ makes the FOC of the inner problem, the one in which $\sqrt{\epsilon}, \theta_2$ and κ_1 are fixed, regular thus allowing us to solve it. This means that just as in previous model, we can concetrate on optimizing the outer problem as the values of parameters $\bar{V_1}, V_2$ and V_3 are uniquely determined by them as the solutions of FOC of the inner problem.

The unfortunate sideeffect of introducing the new group parameter $\bar{V_1}$ into the inner problem is that the outer problem, where we optimize the values of parameters $\sqrt{\epsilon}, \theta_2$ and κ_1 now becomes singular. This is caused by the fact that γ_1, γ_2 and γ_3 are all linear multiples of $\sqrt{\epsilon}$. As such any change in $\sqrt{\epsilon}$ simply translates as scaling in parameters $\bar{V_1}, V_2$ and V_3. This means that changing the value of parameter $\sqrt{\epsilon}$ never brings about a better fit in terms of lowering F. For this reason we pick an arbitrary value of $\sqrt{\epsilon}$ that we use throughout the optimization procedure, e.g. in the simulation studies conducted in the following section we use $\sqrt{\epsilon} = 0.2$. We have only requirement for this value and that is the numerical stability of the whole estimation process. We have found that for some values of $\sqrt{\epsilon}$ the optimization problem is ill behaved in terms of

numerical stability. If we happen to choose a value of $\sqrt{\epsilon}$ for which this is the case we pick a new one until the problem is numericaly stable. This value should also be sufficiently small as the assumption of small ϵ is the basis of the idea behind the asymptotic approximation.

This process allows us to fit the approximation of Fong-Vasicek two factor model very confidently, in terms of cost function F. It does however come with a downside. The group parameters we estimate lose explanatory value. While we can calculate their theoretical values from the parameters used to generate the simulated data, there is no way to go back to the original parameters from the ones we estimate.

5 Model Fit comparison

We choose to evaluate the added value of using the approximation of two factor Fong-Vasicek model over the simpler one factor Vasicek model in two ways. First we conduct a simulation study to compare how well the models we estimate fit yield curves generated using the two factor Fong-Vasicek model. In the second part of this section we look at how well can the models fit real data. In this part we use data about annualized daily interest rates on German zero-coupon bonds, the empirical yield curves, from the period of 10th January 1995 to 30th April 2009.

5.1 Model comparison on simulated data

The point of this section is to assess whether the asymptotic approximation of two factor Fong-Vasicek model, AFVM, is a significant improvement over using the simpler one factor Vasicek model, VM. We measure improvement of fit in terms of the cost function F defined in equation (4.1).

The method we propose for comparing these models is as follows. We choose a reasonably high number of experiments $B = 1000$. We generate B samples of yield curve data that follow the two factor Fong-Vasicek model using the data generating process described in section 3.3. The parameters of the model are taken from Table 3.1. We only present the results for parameter sets 1 and 4 as these contain the most important results. In each sample we employ the algorithms defined in section 4 to estimate the three models that we take into consideration.

For the first model the estimation algorithm is straightforward, the only input needed for the estimation algorithm is a starting value of parameter κ, which we use throughout all experimets $\kappa_0 = 0.1$. In the case of the AFVM the algorithm needs more starting points. The starting values of parameters κ_1 and θ_2 are taken as optimal values of corresponding parameters from the estimated VM. An arbitrary value for the parameter $\sqrt{\epsilon}$ is take to be $\sqrt{\epsilon} = 0.2$ in all samples. We choose a single value for all samples, because changing it for different samples would be needlessly complicated for our purposes. The price for this approach is that in some samples the estimation

algorithm does not reach a stable solution for the last model, or the solution it arrives at makes no sense, e.g. parameters are complex numbers. For the purpose of further study we choose to ignore these samples, noting that there is only a handful, approximately 3%, of them. After omitting these samples we are left with 971 samples that give reasonable results for parameter set 1. For parameter set 4, all the samples were estimated without unreasonable results and we take the full 1000 into consideration.

We compare the quality of the fit by looking at cost function F as defined in (4.1). Figure 5.1 shows the values of the cost function F for our three models with an index of the generated sample on the x axis.

The first thing to notice is that the cost function F for VM is very similar to that of AFVM for parameter set 1. On average the relative improvement of using AFVM over VM, calculated as

$$1 - F_{AFVM}/F_{VM}, \tag{5.1}$$

is 8.18%. It is important to note that using AFVM we arrive at smallest value of F in every observed sample. This should be the case as AFVM has more parameters than VM and should thus be a better fit. Table 5.1 shows a few descriptive statistics about the values of cost function F for the three examined models in parameter set 1.

	Vasicek Model	Approximation of Fong-Vasicek Model
Mean	7.87E-06	7.48E-06
Median	5.88E-06	5.62E-06
Worst fit	6.89E-05	6.84E-05

Table 5.1: Some descriptive statistics of values of cost function F for the two compared models fitted onto data generated using the parameter set 1 from Table 3.1.

The results for AFVM in parameter set 1 seem underwhelming. The improvement over VM, while being present in every considered sample, is only minor. The relative improvement ranges from negligeble, i.e. practically 0, to very significant 70% in the

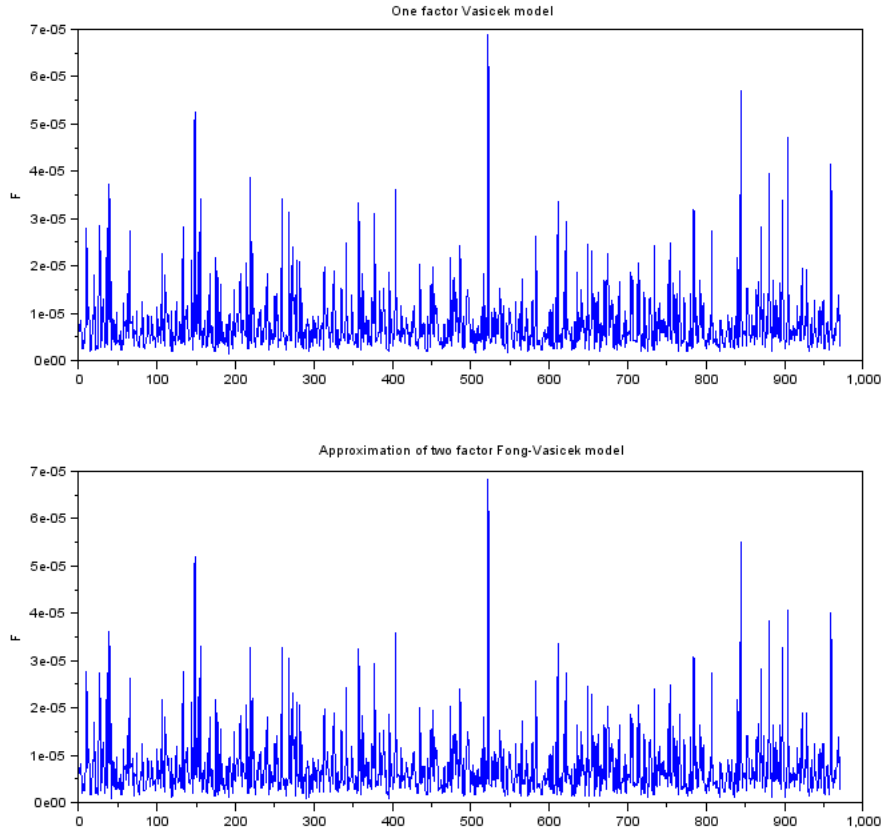


Figure 5.1: The values of cost function F for the estimated fits of one factor Vasicek model and approximation of two factor Fong-Vasicek model in 971 generated samples following the two factor Fong-Vasicek model using the parameter set 1 from Table 3.1.

most extreme case. The average relative improvement is 8.18%, while the median relative improvement is 3.50%.

In order to explain the underwhelming results in parameter set 1 we experimented with changing the parameter values of the generated samples. We found that the comparative performance of VM and AFVM is not significantly impacted by positive correlation ρ using the parameter set 2. The results of the study on this parameter set were almost identical to those of parameter set 1 and their detailed analys is thus omitted.

However, we found that changing the parameter κ_2 changes the results profoundly. Intuitively this should make sense. The unit of fast scale ϵ that defines the accuracy of the asymptotic approximation defining the AFVM is directly linked to the value of

parameter κ_2 by

$$\epsilon = \frac{1}{\kappa_2}.$$

While the fitting procedure we use does not estimate the value of ϵ, remember that its value is fixed, its underlying value clearly plays an important role. Consider Figure 5.2, which shows the values of cost functions F for the two models when we generate the data using parameter set 4. In this figure we can see that the VM and AFVM do not share the scale of the vertical axis anymore. The fit of AFVM in terms of the cost function F is significantly better that that of VM.

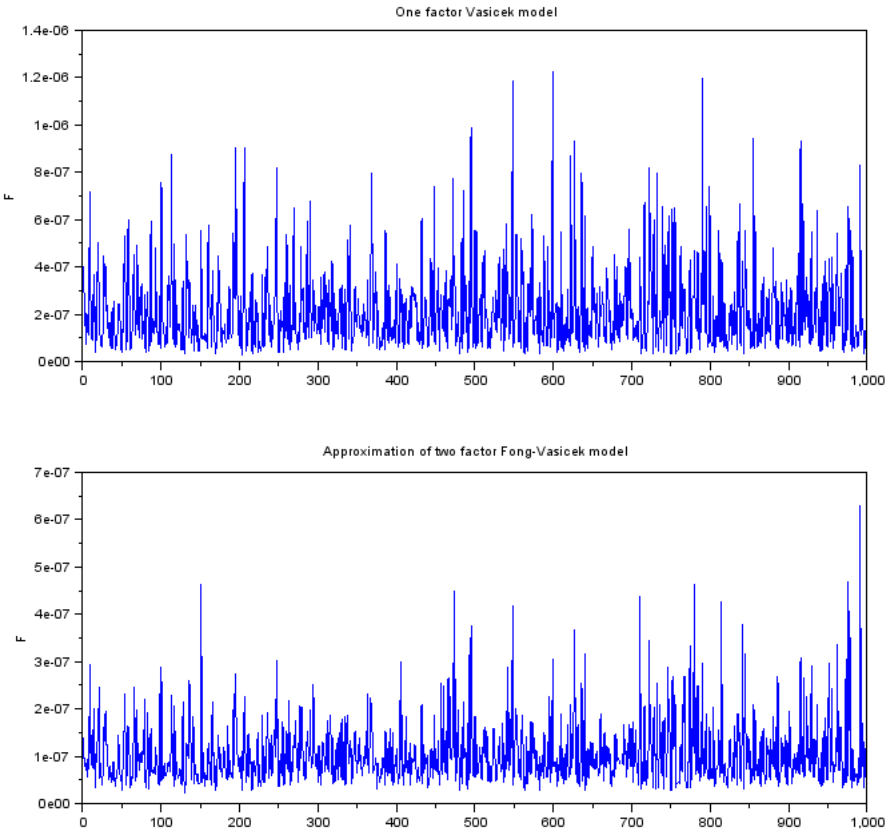


Figure 5.2: The values of cost function F for the estimated fits of constant model, one factor Vasicek model and approximation of two factor Fong-Vasicek model in 1000 generated samples following the two factor Fong-Vasicek model using the parameter set 4 from Table 3.1.

In results on data generated from parameter set 4, the AFVM is still the best approximation every sample. In strong contrast to the results observed using parameter

set 1, the AFVM is in this case a significant improvement. The relative improvement of fit, defined by equation (5.1), averages at 40.36% in samples generated using parameter set 4. The median value of the relative improvement is even higher at 45.23%.

Based on the results of the simulation study we conclude that the AFVM is an improvement over the simpler VM. It outperforms VM in terms of cost function F in every sample generated based on the two factor Fong-Vasicek model with parameter sets 1 through 4 from Table 3.1. The magnitude of the improvement is closely tied with the value of parameter κ_2. With the increasing value of parameter κ_2 the improvement that AFVM brings over VM increases as well. This is because it leads to faster scaling volatility making AFVM a more accurate approximation of the two factor Fong-Vasicek model.

5.2 Model comparison on real data

This section presents the results of fitting VM and AFVM onto real data. The data we are fitting in this section consists of daily yield curves implied by German bonds in the period from January 10th 1995 to April 30th 2009. The observations do not include weekends and there are sometimes days missing from the week as well. Overall we have data about 3735 observed yield curves. To illustrate this data we refer the reader to Figure 1.1, which contains the development of the rate on 3 month bonds, the shortest maturity available.

To compare the estimation techniques derived in section 4 on more than a single sample we divide our observations into blocks. We define periods 1 through 10 as blocks of 250 observation. Period 1 is the most recent, thus ending with the last day of available data, i.e. April 30th 2009. Period 2 is defined as 250 observations before the first observation of period 1, and so on. For each of these periods we fit the VM and AFVM. The optimal value of cost function F for all periods is contained the first two columns of Table 5.2. The last column of this table contains the relative improvement of fit, defined by equation (5.1).

We see that the AFVM is not a vast improvement over the simpler VM. As was

Values of cost function F	One factor Vasicek model	Approximation of two factor Fong-Vasicek model	Relative Improvement
Period 1	8.88E-08	8.28E-08	6.754%
Period 2	3.05E-08	3.05E-08	0.008%
Period 3	3.72E-08	3.70E-08	0.413%
Period 4	1.31E-07	1.31E-07	0.001%
Period 5	1.25E-08	1.25E-08	0.271%
Period 6	3.27E-08	2.39E-08	26.858%
Period 7	5.62E-08	5.61E-08	0.301%
Period 8	4.39E-08	4.22E-08	3.819%
Period 9	6.54E-08	6.41E-08	1.932%
Period 10	2.17E-07	2.12E-07	2.519%

Table 5.2: The optimal values of cost function F for the fits of the Vasicek model and the approximated Fong-Vasicek model. Each row represents a 250 day period.

the case in simulated data, the fit is improved in each period by estimating AFVM in comparison to estimating VM. However the added value of the more complex model seems only minimal. The only period in which the improvent seems significant is period 6 with improvent of almost 27%. On the other hand in half of the periods the improvement is less than 1%. Figure 5.3 illustrates the fitted models on an example of 6 yield curves from period 3. For a better comparison of numerical values we provide Table 5.3. This table contains the numerical values for two yield curves, number 3019 and 3185, both included in Figure 5.3. One thing we noticed is that the fitted AFVM offeres a much larger variation in the shapes of possible yield curves. Notice that the dashed lines, representing the fitted VM, are almost exactly the same in terms of shape. In contrast the solid lines representing the fitted AFVM produce a much wider range of shapes, which seems, at least visually, to be closer to the shape of the observed data. However as evidenced by the yield curve number 3185 sometimes both models seem far

from the observations. The result represented by yield curve number 3185 is perhaps the most frequent one seen in the data. The models both seem to not completely grasp the underlying process and as a result they end up far further from the observations.

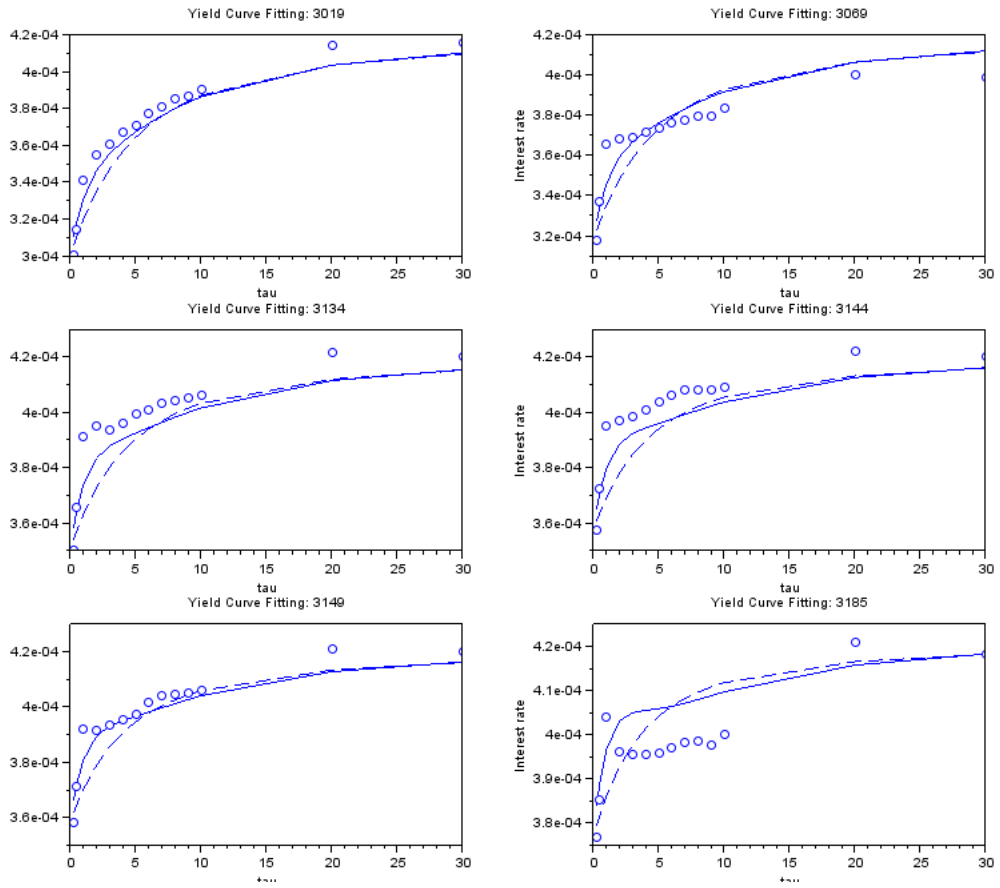


Figure 5.3: The resulting fits of the Vasicek model and the approximation of Fong-Vasicek model on 6 yield curves from period 3. The data are represented by dots, the fitted Vasicek model is represend by a dashed a line and the fitted approximation of Fong-Vasicek model is represented by a solid line.

The results however are not only negative. The optimal values of cost function F are very low, meaning that the models are actually a pretty good fit for the data even though the visual evidence does not seem to indicate so. We conclude that these models, while they are not perfect are at least reasonable representations of the data we analyzed.

We propose, that the fact that AFVM is not a significant improvement over the VM can be explained by a relatively large value of ϵ, the unit of fast scaling volatility. We

τ	O (3019)	VM (3019)	AFVM (3019)	O (3185)	VM (3185)	AFVM (3185)
0.25	0.0300%	0.0306%	0.0310%	0.0377%	0.0379%	0.0384%
0.5	0.0314%	0.0311%	0.0319%	0.0385%	0.0382%	0.0389%
1	0.0341%	0.0320%	0.0331%	0.0404%	0.0386%	0.0397%
2	0.0355%	0.0335%	0.0346%	0.0396%	0.0393%	0.0403%
3	0.0360%	0.0347%	0.0355%	0.0395%	0.0398%	0.0405%
4	0.0367%	0.0357%	0.0362%	0.0395%	0.0401%	0.0406%
5	0.0371%	0.0364%	0.0367%	0.0396%	0.0404%	0.0406%
6	0.0378%	0.0371%	0.0372%	0.0397%	0.0406%	0.0406%
7	0.0381%	0.0376%	0.0376%	0.0398%	0.0408%	0.0407%
8	0.0385%	0.0380%	0.0380%	0.0398%	0.0410%	0.0408%
9	0.0387%	0.0384%	0.0383%	0.0398%	0.0411%	0.0409%
10	0.0391%	0.0387%	0.0386%	0.0400%	0.0412%	0.0410%
20	0.0415%	0.0404%	0.0404%	0.0421%	0.0417%	0.0416%
30	0.0416%	0.0410%	0.0410%	0.0418%	0.0418%	0.0418%

Table 5.3: The numerical values of observed yield curves, denoted by O, number 3019 and 3185 along with their estimates by VM and AFVM.

have observed similar results in the simulated samples generated using parameter set 1 in the previous section. If the unit of the fast scale is not small enough the AFVM does not a add substential amount of insight into the workings of the underlying process governing the short rate and the yield curves.

More work could be done in order to determine a priori whether the sample exhibits the evidence of fast scaling volatility. A screening method could then be devised which would allow us to assess whether an asymptotic approximation using fast scaling volatility brings a significant benefit over using other models. A step in this direction is hinted at in [9], where the authors mention using variograms to determine the period of mean-reversion in the unobserved volatility process. If this period of mean-reversion is deemed sufficiently short, the assumption of fast scaling volatility should hold.

Conclusion

This diploma thesis focused on the models of the short rate. We introduced the one factor Vasicek model and the two factor Fong-Vasicek model with stochastic volatility. An asymptotic approximation of Fong-Vasicek model was also described as a middle ground between these two models. The theoretical usefulness of this approximations is that it allows for more shapes of the yield curves than the one factor model. Moreover it has less parameters than the two factor model, and perhaps most importantly, the pricing function does not depend on the value of unobservable volatility. These properties make this approximation very attractive from a modeling point of view.

The main goal of this diploma thesis was to come up with an algorithm that estimates the parameters of this approximation. This algorithm was defined in the fourth section. We have also successfully tested the validity of this algorithm by using it to fit randomly generated yield curves with known parameters. Because the nature of the algorithm does not allow us to directly compare the estimated parameters, we define a different measure of success using a cost function F defined in equation (4.1). This cost function punishes the difference between a yield curve proposed by the fitted model and the observed data.

We find that both Vasicek model and the asymptotic approximation of Fong-Vasicek model are reasonable models for data generated following the Fong-Vasicek model. The added value of the approximation depends strongly on the underlying value of the unit of fast scaling volatility ϵ. We illustrate this in the last section by comparing resulting optimal values of the cost function F for two different parameter sets where the main difference is in the parameter ϵ.

We also find that both models are reasonable when fitting the real world data about German yield curves. In terms of the optimal values of cost function F the fit obtained in real data is comparable to that observed in the generated data, which we consider a very good result. Dividing our sample into blocks we find that the added value of the approximation is not very large in many blocks of observations. From a qualitative point of view however, we observe that the shapes of yield curves offered by the

approximation are much more diverse than those offered by the Vasicek model. Furthermore the approximation was a superior model in every observed sample in terms of the smaller optimal value of cost function F. We thus conclude that the aims of this diploma thesis were met. Our estimation procedures fits the approximation of Fong-Vasicek model with stochastic volatility onto yield curve data and has very encouraging results in doing so for both generated and real world data.

References

[1] Andersen, T. G., Lund, J.: *Estimating continuous-time stochastic volatility models of short-term interest rate*, Journal of Econometrics, **77**, pages 343-377, 1997.

[2] Black, F., Scholes, M.: *The Valuation of Option Contracts and Test of Market Efficiency*, Journal of Political Economy, **27**, pages 399-417, 1972.

[3] Black, F., Scholes, M.: *The pricing of options and corporate liabilities*, Journal of Political Economy, **81**, pages 637-654, 1973.

[4] Brenner, R. J., Harjes, R.H., Kroner, K. F.: *Another Look at Models of the Short Term Interest Rate*, Journal of Financial and Quantitative Analysis, **35**, pages 191-215, 2000.

[5] Brigo, D., Mercurio, F.: *Interest rate models - Theory and practice. With smile, inflation and credit*, Springer, 2006.

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

[7] Chattová, S.: *Kalibrácia konvergenčného modelu úrokových mier Vašíčkovho typu*, Diploma thesis, , FMFI, UK, 2013.

[8] Cotton, P.: *An analytic approach to Ornestein-Uhlenbeck processes with fluctuating parameters and applications in the modeling of fixed income securities*, PhD thesis, Stanford University, 2001.

[9] Cotton, P., Fouque, J. P., Papanicolaou, G., Sircar, R.: *Stochastic Volatility Corrections for Interest Rate Derivatives*, 2003, available on the internet (20.4.2014): `http://www.princeton.edu/~sircar/Public/ARTICLES/cfps_www.pdf`

[10] Danilov, D., Mandal, P. K.: *Estimation of the volatility component in two factor stochastic volatility short rate models*, EURANDOM reports, **40**, 2000, available on the internet (20.4.2014):
`http://www.eurandom.tue.nl/reports/2000/040-report.pdf`

[11] Danilov, D., Mandal, P. K.: *Cross sectional efficient estimation of stochastic volatility short rate models*, EURANDOM reports, 36, 2001, available on the internet (20.4.2014)
`http://www.eurandom.tue.nl/reports/2001/036-report.pdf`

[12] Derman, E., Kani, I.: *The Volatility Smile and Its Implied Tree*, RISK, pages 139-145, 1994.

[13] Fong, H. G., Vasicek, O. A.: *Fixed-Income Volatility Management*, Journal of Portfolio Management,**17**, pages 41-46, 1991.

[14] Fouque, J.P, Papanicolaou, G., Sircar, R.: *Mean-Reverting Stochastic Volatility*, International Journal of Theoretical and Applied Finance, **3**, pages 101-142, 2000.

[15] Kwok, Y. K.: *Mathematical Models of Financial Derivatives*, Springer-Verlag, 1998.

[16] Longstaff, F., Schwartz, E.: *Interest rate volatility and the term structure: a two factor general equilibrium model*, Journal of Finance, pages 1259-1282, **47**, 1992.

[17] Mandelbrot, B. B.: *The Variation of Certain Speculative Prices*, The Journal of Business, **4**, pages 394-419, 1963.

[18] Melicherčík, I., Olšarová, L., Úradníček, V.: *Kapitoly z finančnej matematiky*, Vydavateľstvo EPOS, Bratislava, 2005.

[19] Rebonato, R.: *Interest Rate Option Models Inderstanding Analysing and Using Models for Exotic Interest Rate Options*, Chichester, Wiley, 1996.

[20] Selečéniová, R.: *Rýchla časová škála volatility vo Fong-Vašíčkovom modeli*, Diploma thesis, , FMFI, UK, 2012.

[21] Ševčovič, D., Urbánova Csajková, A.: *On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model*, Central European J. of Operation Research, **13**, pages 169-188, 2005.

[22] Ševčovič, D., Stehlíková, B., Mikula K.: *Analytické a numerické metódy oceňovania finančných derivátov.*, Nakladateľstvo STU, Bratislava, 2009.

[23] Stehlíková, B., Ševčovič, D.: *On a volatility averaging in a two-factor interest rate model*, Algoritmy, pages 325-333, 2005.

[24] Stehlíková, B.: *Mathematical analysis of term structure models*, Dissertation Thesis, FMFI, UK, 2008.

[25] Stehlíková, B., Ševčovič, D.: *On the singular limit of solutions to the Cox-Ingersoll-Ross interest rate model with stochastic volatility*, Kybernetika, **40**, 2008.

[26] Suli, E., Mayers, D.: *An Introduction to Numerical Analysis*, Cambridge University Press, 2003.

[27] Vasicek, O. A.: *An Equilibrium Characterization of the Term Structure*, Journal of Financial Economics, **5**, pages 177-188, 1977.