Abstract

KEYWORDS: Arbitrage Free model, Equilibrium Models, Log-normal model, HJM Model, Ornstein-Uhlenbeck Process

Dedication

This work is dedicated to the FOGNE'S Family for their endless love, support, encouragement and most especially to my beloved sister SHANELLE.

Acknowledgment

First and foremost, I praise and thank the Lord almighty, for his showers of blessings throughout my life.

I would like to thank my advisor, Dr. Vajira Manathunga, for his continuous support throughout my master's study and research, for his patience, motivation, understanding and enthusiasm.

I would also like to give special thank to Prof. Jator Samuel. for his advices, support and contribution during this work.

Finally, I must express my very profound gratitude to my Mum for providing me with unfailing support and continuous encouragement throughout my years of study. This accomplishment would not have been possible without her. Thank you.

Contents

Chapter ONE

Introduction

One of the most computationally challenging part of modern financial theory nowadays is the interest rate and more particularly its dynamics. The real challenge in modeling these interest rates arises due to the existence of a term structure of interest rates inbuilt in the shape of the forward curve. The correct modeling of the term structure of interest rate (the stochastic behavior) over time is crucial for the construction of reliable and realistic models for interest rate derivatives. An interest rate model is a probabilistic description of the future evolution of interest rates. Based on current information, future interest rates are uncertain: An interest rate model is a characterization of that uncertainty [pdf12]. For financial instrument valuation and risk estimation, one wants to use models that are arbitrage free and matched to the currently observed term structure of interest rates. *"Arbitrage free"* means just that if one values the same cash flows into different ways, one should get the same result[pdf12]. An arbitrage free model will produce the same value for the structure viewed either ways. This is also known as the *law of one price*. Available are a good number of term structure models used to stimulate the motion of interest rates. Understanding the characteristics of different term structure models can help the user choose the most appropriate model for his/her application. One very important step in using a term structure model relies in the users ability to choose appropriate parameters. Even if the form of the term structure model incorporates all the desirable historical movements in yields (such as mean reversion), such models will provide poor results if the parameters are carelessly chosen [pdf3].

The aim of this paper is to provide a review of the modeling techniques of the term structure applicable to default free bonds and other interest rate derivatives. We discuss term structure modeling only in its crudest outline and focus more on the short rate models, in which the stochastic state variable is taken to be the instantaneous short rate.

This paper is outlined as follows: Section 1 introduces...; Section 2...;

1.1 Research Purpose

1.2 Research Limitations

Chapter TWO

Preliminaries

The purpose of this chapter is to fix some terminologies that will be used throughout this work, and to present some few tools necessary for modeling the stochastic behavior of interest rates. The statement of these terminologies and tools can be found in [18,35,36,37].

Definition 2.0.1. *A **Bond** is a financial debt instrument requiring the issuer(borrower) to repay to the lender(investor) the amount borrowed plus interest over a specific period of time(maturity).*

The size of the bond market (debt or fixed income market) in 2009 was an estimated of $82.2 trillion of which the size of the USA bond market debt was $31.2 trillion according to "Bank of International Settlement" [33]. Bond prices are less volatile than stock price, which makes them safer investments. However, they also have some risks, such as the *default risk* and the *interest rate risk*. For the purpose of this rate, we would be concerned with the interest rate risk.

Definition 2.0.2. *A **Coupon Bond** is a contract which guarantees the holder a payment stream during the life of the bond. This instrument have a common property, that it provides the holder with a deterministic cash flow, and for this reason bonds are also known as Fixed Income Instruments.*

Definition 2.0.3. *A **Zero Coupon Bond** with maturity date T is a contract which guarantees the holder 1 unit of money to be paid at the maturity date T. The price at time t of a bond with maturity date T is denoted by $P(t,T)$.*

The bond price $P(t,T)$ is a stochastic object with two variables, t and T, and for each outcome in the underlying sample space, the dependence upon these variables is very different. For a fixed value of t, $P(t,T)$ is a function of fixed maturity T that could be viewed as a stochastic process which gives the prices at different times with very irregular trajectory. The graph of this function is called the **Term Structure** or **yield curve** at time t. That is, it is the relationship between interest rates or bond yields and different terms (or maturities). Below are the three hypothetical yield curves;

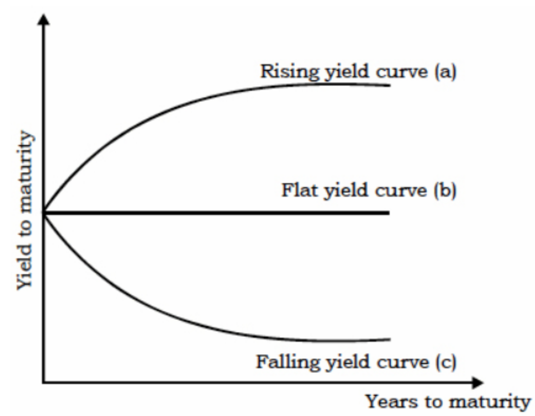


Figure 2.1: Hypothetical Yield Curves

The figure above shows three shapes that have appeared with some frequency over time when the term structure—the yield-to-maturity, or the spot rate, at successive maturities against maturity is plotted. Curve (a) shows an upward-sloping yield curve; that is, yield rises steadily as maturity increases. This shape is commonly referred as a normal or upward-sloping yield curve. Curve (c) shows a downward-sloping or inverted yield curve, where yields decline as maturity increases. Curve (b) shows a flat yield curve.

Two major theories have emerged to account for these shapes: the **expectations theory** and the **market-segmentation theory**.

There are three forms of the expectations theory: the **pure expectations theory**, the **liquidity theory**, and the **preferred-habitat theory**.

The Pure Expectation Theory

According to this theory, the forward rate exclusively represent expected future rates. Thus the entire term structure at a given time reflects the market's current expectations of the future short-term rates. Under this theory, a rising term structure, as shown in curve (a), must indicate that the market expects short-term rates to rise throughout the relevant future. Similarly, a flat term structure reflects an expectation that the future short-term rates will be mostly constant, and a falling term structure must reflect an expectation that future short term rates will decline steadily. The drawback of this theory is that it does not account for the risks associated with investing in bonds.

The Liquidity Theory

There is risk in holding a long term bond for one period, and that risk increases with the bond's maturity because maturity and price volatility are directly related.

Given this uncertainty, and the fact that investors do not like uncertainty, some economists and financial analysts have suggested a different theory (The Liquidity Theory). This theory states that investors will hold longer term maturities if they are offered a long-term rate higher than the average of expected future rates by a risk premium that is positively related to the term to maturity. That is, forward rates should reflect both interest rate expectations and liquidity premium (which is really a risk premium), and the premium should be higher for longer maturities.

The Preferred-Habitat Theory

This theory also adopts the view that the term structure reflects the expectation of the future path of interest rates as well as a risk premium. However, it rejects the the assertion that the risk premium must rise uniformly with maturity. Thus, it proposes that the shape of the yield curve is determined by both expectations of future interest rates and a risk premium, positive or negative, to induce market participants to shift out of their preferred habitat.

Market-Segmentation Theory

This theory recognizes that investors have preferred habitats dictated by the nature of their liabilities. It also proposes that the major reason for the shape of the yield curve lies in asset/liability management constraints (either regulatory or self-imposed) and creditors (borrowers) restricting their lending to specific maturity sectors. This theory is however different from the preferred-habitat theory in that it assumes that neither investors nor borrowers are willing to shift from one maturity sector to another to take advantage opportunities arising from differences between expectations and forward rates. Thus, for this theory, the shape of the yield curve is determined by supply of and demand for securities within each maturity sector.

Definition 2.0.4. *The **Principal Value** (Par-Value, Face Value, Redemption Value or Maturity Value) is the stated face value of the bond which is paid at the time of maturity.*

Definition 2.0.5. *A **discount bond** is a bond that is issued at a lower price than its par value or a bond that is trading in the secondary market at a price that is below the par value. It is similar to a zero coupon bond, only that the latter does not pay interest. A bond is considered to trade at discount when its coupon rate is lower than the prevailing interest rates.*

We assume the following three conditions in this work

(i) There exist a frictionless market for bonds with each maturity. That is, no tax, no transaction cost, equal availability of information to all traders etc.

(ii) The relation $P(t,t) = 1$ for all t.

(iii) For each fixed t, the bond price $P(t,T)$ is differentiable with respect to time of maturity (assuming T can be taken as large as possible).

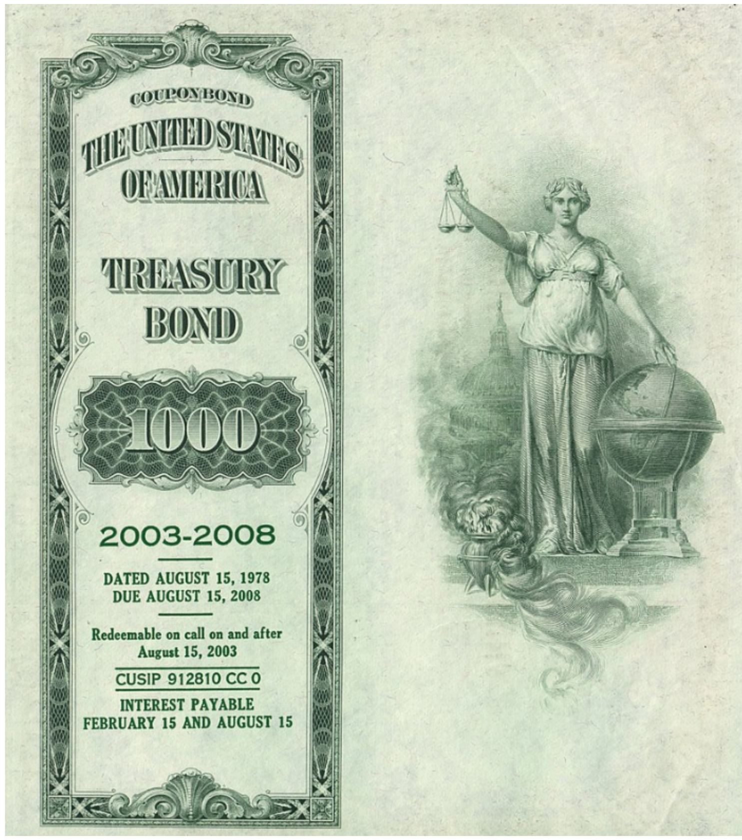


Figure 2.2: 1978 $1000 U.S. Treasury Bond

Definition 2.0.6. *The* ***Yield*** *(Yield to Maturity) is the interest rate that the bond holder earns for buying the bond. The issuer of the bond will make the interest payment according to the specified yield.*

Definition 2.0.7. ***Arbitrage*** *is the type of transaction in which an investor wants to make to make profit when the same assets are trading at two different prices. That is, the transaction here involves the investor buying the assets at lower prices and selling them at higher prices. This is often referred to as* ***free lunch****.*

We note that the assumption that $P(t,t) = 1$ should always hold in order to avoid arbitrage because suppose that $P(t,t) < 1$, then this means that the price of a zero-coupon bond at time t which expires on the same day is smaller than 1. Therefore, the arbitrageurs would buy this bond and on the same day the bond expires they get 1 unit of money.

Definition 2.0.8. *Any asset which has strictly positive prices for all* $t \in [0,T]$ *is called a* ***numeraire****.*

Definition 2.0.9. *The prices* Z_n *of other asset denominated in terms of numeraire* Z_1 *are called* ***relative prices*** *and denoted by* $Z_n^* = \frac{Z_n}{Z_1}$.

Another assumption we lay is that there exists a trade-able asset called **money-market account**, defined as the stochastic process satisfying

$$dD_t = r_t D_t dt, \qquad D_0 = 1 \tag{2.1}$$

Thus, cash can be invested in an asset that is risk-less over short periods of time, an accrues interest at the short rate r_t. Therefor, we have that

$$D_t = e^{\int_0^t r_s ds} \tag{2.2}$$

Definition 2.0.10. *A **risk premium** or **spread** is the minimum amount of money by which the expected return on a risky asset must exceed the known return on a risk-free asset in order to induce an individual to hold the risky asset rather than the risk-free asset.*

Below are some commonly used interest rates in financial mathematics

2.1 Basic Rates and Notation

Definition 2.1.1. *The simple or simply compounded **spot rate** for the period $[t, T]$ prevailing at time t is defined as*

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

so that a bond price expressed in terms of the spot rate is given as

$$P(t, T) = \frac{1}{(1 + r(t, T)(T - t))} \tag{2.4}$$

Definition 2.1.2. *The simply compounded **forward rate** for the period $[T_1, T]$ prevailing at time t is defined as*

$$f(t; T_1, T) = \frac{P(t, T_1) - P(t, T)}{(T - T_1)P(t, T)} \tag{2.5}$$

Definition 2.1.3. *The continuously compounded **forward rate** for the period $[T_1, T]$ prevailing at time t is defined as*

$$f(t; T_1, T) = -\frac{\ln P(t, T) - P(t, T_1)}{(T - T_1)} \tag{2.6}$$

Definition 2.1.4. *The continuously compounded **spot rate** for the period $[t, T]$ prevailing at time t is defined as*

$$y(t, T) = f(t; t, T) = -\frac{\ln P(t, T)}{(T - t)} \tag{2.7}$$

Definition 2.1.5. *The **instantaneous forward rate** defined as*

$$f(t, T) = \lim_{T_1 \to T} f(t; T_1, T) = -\frac{\ln P(t; T)}{\partial T} \tag{2.8}$$

Definition 2.1.6. *The **instantaneous short rate** defined as*

$$r(t) = f(t,t) = \lim_{(T_1 - T)\to 0} y(t,T) \tag{2.9}$$

The short rate $r(t)$ is the short that applies to an infinitesimally short period of time at time t. Short rate is actually a theoretical entity. It does not exist in real life and can not be observed directly.
We note that equation (2.6) together with the condition that $P(T,T) = 1$ gives

$$P(t,T) = e^{-\int_t^T f(t,s)ds} \tag{2.10}$$

For an uncertain future, instantaneous rates can be considered as random process between times t and $t + dt$ and then we can define the price of a zero coupon bond in an uncertain future based on stochastic calculus (Brownian motion).

2.2 Basic concepts of stochastic calculus.

In this section we briefly present the basic definitions and theorems of stochastic calculus which will be needed to formulate the models considered in this work.

All our models will be set up in a given complete probability space $(\Omega, \mathcal{F}_t, \mathcal{P})$ and an augmented filtration $\mathcal{F}_t$ generated by a standard Brownian motion in $\mathbb{R}$

Definition 2.2.1. *A **filtration** $(\mathcal{F}_t)_{t\geq 0}$ of a probability space $(\Omega, \mathcal{F}_t, \mathcal{P})$ is a family of σ-algebra(a family of events including the empty set that is closed under complementation and countable unions) $\mathcal{F}$ indexed by $t \in [0,\infty]$, all contained in $\mathcal{F}$, satisfying*

(i) if $s \leq t$ then $\mathcal{F}_s \subset \mathcal{F}_t$, and

(ii) $\mathcal{F}_\infty = \sigma(\cup_{t\geq 0}\mathcal{F}_t)$

$(\mathcal{F}_t)_{t\geq 0}$ is used to model the flow if information.

Definition 2.2.2. *The **stochastic process** is a parametrized collection of random variables $\{X_t\}_{t\in T}$ defined on a probability space $(\Omega, \mathcal{F}_t, \mathcal{P})$ and assuming values in $\mathbb{R}$.*

Definition 2.2.3. *A stochastic process $(X_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}_t, \mathcal{P})$ is said to be **adapted** to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if for each $t \geq 0$ the random variable X_t is measurable relative to $(\mathcal{F}_t)$. (A random variable Y is measurable with respect to a $\sigma-$algebra if for every open set U the event $Y \in U$ is in the $\sigma-$algebra.)*

Definition 2.2.4. *Let $\mathcal{X} = \{x_1, \ldots, x_N\}$ be the set of a finite number of discrete states. The stochastic process $X = \{X_t, t \in \mathbb{R}^+\}$ is a **continuous time Markov chain** if it satisfies the following Markov property:*

$$P(X_t \in \mathcal{B}|X_s = x) = P(X_t \in \mathcal{B}|X_{r_1} = x_1, \ldots, X_{r_n} = x_n, X_s = x) \tag{2.11}$$

for all Borel subsets $\mathcal{B} \subseteq \mathbb{R}$ for $0 \leq r_1 \leq \ldots \leq r_n < s < t$ and for all $x_1, \ldots, x_n, x \in \mathbb{R}$ for which the conditional probabilities are defined.

For fixed s, x and t the transition probability $P(X_t \in \mathcal{B}|X_s = x)$ is a probability measure on the $\sigma-$algebra $\mathcal{A}$ of Borel subsets of $\mathbb{R}$ such that

$$P(X_t \in \mathcal{B}|X_s = x) = \int_{\mathcal{B}} p(s, x; t, y)dy \tag{2.12}$$

for all $\mathcal{B} \in \mathcal{A}$. The quantity $p(s, x; t, \cdot)$ is the transition density which play the role of a transition matrix in Markov chains.

Definition 2.2.5. *A **Markov process** $X = \{X_t, t \in \mathbb{R}^+\}$ is a **diffusion process** if the following limits exists for all $\epsilon > 0$, $s \geq 0$ and $x \in \mathbb{R}$:*

$$\lim_{t\to s} \frac{1}{t-s} \int_{|y-x|>\epsilon} p(s, x; t, y)dy = 0, \tag{2.13}$$

$$\lim_{t\to s} \frac{1}{t-s} \int_{|y-x|<\epsilon} (y-x)p(s, x; t, y)dy = \alpha(s, x), \tag{2.14}$$

$$\lim_{t\to s} \frac{1}{t-s} \int_{|y-x|<\epsilon} (y-x)^2 p(s, x; t, y)dy = \beta^2(s, x) \tag{2.15}$$

where $\alpha(s, x)$ is the *drift* and $\beta^2(s, x)$ *diffusion coefficient* at time s and position x. The simplest and most fundamental diffusion process is Brownian motion which is sometimes called the Wiener process.

Definition 2.2.6. *A stochastic process $\{X(t), t \geq 0\}$ is called a **Wiener process** (also called Brownian motion), if it satisfies the following properties:*

(i) X(0) = 0 with probability 1;

(ii) every increment X(t+$\triangle$t) - X(t) has the normal distribution N(0,$\triangle$t);

(iii) the increments $X(t_n) - X(t_{n-1}), X(t_{n-1}) - X(t_{n-2}), \cdots, X(t_2) - X(t_1)$ for $0 \leq t_1 < \cdots < t_n$ are independent.

Definition 2.2.7. *An **Ito process** or **stochastic integral** is a stochastic process $(\Omega, \mathcal{F}_t, \mathcal{P})$ adopted to $\mathcal{F}_t$ which can be written in the form*

$$X_t = X_0 + \int_0^t U_s ds + \int_0^t V_s dB_s, \tag{2.16}$$

where $U, V \in \mathcal{L}_2$. As a shorthand notation, we will write (2.1) as

$$dX_t = U_t dt + V_t dB_t.$$

Theorem 2.2.1. ***(Ito formula)*** *Let X_t be an Ito process $dX_t = U_t dt + V_t dB_t$. Suppose $g(x) \in C^2(\mathbb{R})$ is a twice continuously differentiable function (in particular all second partial derivatives are continuous functions). Suppose $g(X_t) \in \mathcal{L}_2$. Then $Y_t = g(X_t)$ is again an Ito process and*

$$dY_t = \frac{\partial g}{\partial x}(X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t)(dX_t)^2$$

Using the notation convention for $dX_t = U_t dt + V_t dB_t$ and $(dX_t)^2$, we can rewrite the Ito formula as

$$dY_t = \left(\frac{\partial g}{\partial x}(X_t)U_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t)V_t^2\right)dt + \frac{\partial g}{\partial x}(X_t)V_t dB_t \tag{2.17}$$

Definition 2.2.8. *Give the probability space* $(\Omega, \mathcal{F}_t, \mathcal{P})$ *and the filtration* $(\mathcal{F}_t)_{t\geq 0}$. *An adapted family* $(M_t)_{t\geq 0}$ *of integrable random variables (that is,* $E(|M_t|) < +\infty$ *) for any* t *is :*

(i) a ***martingale*** *if, for any* $s \leq t$, $E(M_t|\mathcal{F}_s) = M_s$,

(ii) a ***supermartingale*** *if, for any* $s \leq t$, $E(M_t|\mathcal{F}_s) \leq M_s$,

(iii) a ***submartingale*** *if, for any* $s \leq t$, $E(M_t|\mathcal{F}_s) \geq M_s$.

Definition 2.2.9. *Two probability measures* $\mathcal{P}$ *and* $\mathcal{Q}$ *are said to be equivalent if for every event* A, $\mathcal{P}(A) = 0$ *if and only if* $\mathcal{Q} = 0$.

Example 2.1: Let Z be a random variable such that $E[Z] = 1$ and $Z > 0$. Define a new measure $\mathcal{Q}$ by

$$\mathcal{Q}(A) = E[Z]_{1_A} = \int_A Z d\mathcal{P} \tag{2.18}$$

for every event A. The $\mathcal{A}$ and $\mathcal{Q}$ are equivalent.

Theorem 2.2.2. *(**Radon-Nikodym**). Two measures* $\mathcal{P}$ *and* $\mathcal{Q}$ *are equivalent if and only if there exists a random variable* Z *such that* $E[Z] = 1$, $Z > 0$ *and* $\mathcal{Q}$ *is given by (2.13).*

Definition 2.2.10. *Let* $(\Omega, \mathcal{F}_t, \mathcal{P})$ *be a probability space and* $\mathcal{Q}$ *be any probability measure satisfying the following:*

(i) $\mathcal{Q}$ *is equivalent to* $\mathcal{P}$ *(that is, both measures have the same negligible sets);*

(i) the relative price process Z_n^* *are martingales under* $\mathcal{Q}$ $\forall n$ *(that is, for* $s \leq t$ *we have* $E^{\mathcal{Q}}[Z_n^*(t)|\mathcal{F}_s] = Z_n^*(s)$*)*

Then the measure $\mathcal{Q}$ *are called* ***equivalent martingale measures (risk-neutral measure)****.That is the probability distribution for an asset transformed so that the expected return on the asset is the risk-free rate.*

Theorem 2.2.3. *(**Unique Equivalent Martingale Measure**) A continuous time economy is arbitrage free and every security is attainable if for every choice of numeraire* $\exists$ *a unique equivalent martingale measure.*

We note that for different kind of numeraire $\exists$ different unique equivalent martingale measures. The numeraire we consider here is the value of a bank account since it is assumed to earn a constant interest rate r and since it is strictly positive.

Definition 2.2.11. *An* ***Ornstein-Uhlenbeck Process*** *is the unique solution of the following equation*

$$dX_t = -cX_t dt + \sigma d\Phi_t$$

$$X_0 = x$$

where c and σ. It can be explicitly written as

$$X_t = xe^{-ct} + \sigma e^{-ct} \int_0^t e^{-cs} d\Phi_s \tag{2.19}$$

Now that we have defined some basic financial and stochastic notions, let's now start looking stochastic one factor interest rate models and if possible their explicit solutions.

Chapter THREE

Short Term Interest Rate Models

Short rate models define the future trend of interest rates through the description of the future trend of the short rate. The short rate here $r(t)$ is the interest rate at which an entity can borrow money for a short period of time from time t [6]. The current short rate does not specify the entire yield curve. However, no-arbitrage arguments show that under some fairly relaxed technical conditions of $r(t)$ as a stochastic process under the a risk-neutral measure $\mathcal{Q}$, the price at time t of a zero-coupon bond maturing at time T is given by

$$P(t,T) = E\left[e^{\left(-\int_t^T r(s)ds\right)} \Big| \mathcal{F}_t\right] \tag{3.1}$$

where $\mathcal{F}_t$ is the natural filtration for the process, thus specifying a model for the short rate specifies the future bond prices. To specify the structure of interest rate, we begin by assuming that it follows a diffusion process:

$$dr = \mu(r,t)dt + \tau(t,T)d\Phi(t) \tag{3.2}$$

where $\mu(r,t)$ is the deterministic component, $\tau(t,T)$ explains the randomness of the process and Φ is a Wiener process. Based on this assumption, we derive the fundamental pricing equation from which the price of a pure discount bond.

The Long/Short Portfolio

Suppose we have a portfolio made of two zero-coupon bonds, say a long one unit of a short term bond with maturity date T_1 and a short $\triangle$ units of a long-term bond with maturity date $T_2 > T_1$. Suppose the two prices for the maturity dates T_1 and T_2 are $P_1(r,t)$ and $P_2(r,t)$ at time t when the current interest rate is r. Let's denote the value of this portfolio by Π. So the value of the portfolio Π at time t is given to be

$$\Pi(r,t) = P_1(r,t) - \triangle P_2(r,t) \tag{3.3}$$

In order to make this portfolio insensitive to changes in interest rates , that is risk-less, we choose $\triangle$ so that the sensitivity of the portfolio Π to changes in interest rate r_t is equal to zero [7]:

$$\frac{\partial \Pi(r,t)}{\partial r} = 0 \tag{3.4}$$

It follows from equation (3.3) that

$$\frac{\partial P_1(r,t)}{\partial r} - \triangle \frac{\partial P_2(r,t)}{\partial r} = 0 \tag{3.5}$$

From equation (3.5), we obtain the hedging strategy:

$$\triangle = \frac{\partial P_1/\partial r}{\partial P_2/\partial r} \tag{3.6}$$

This hedging strategy is inherent in the sense that if we are long the bond with maturity T_1, then we must short the bond with maturity T_2 according to the relative sensitivity of each bond to changes in interest rates [7].

The process of the "hedged" portfolio between t and $t + dt$ is given by Ito's lemma to be

$$d\Pi_t = \{(\frac{\partial \Pi}{\partial t}) + (\frac{\partial \Pi}{\partial r})\eta(\xi - r_(t)) + \frac{1}{2}(\frac{\partial^2 \Pi}{\partial r^2})\tau^2\}dt + (\frac{\partial \Pi}{\partial r})\tau d\Phi_t \tag{3.7}$$

By the use of equation (3.4), equation (3.7) then becomes

$$d\Pi_t = \{(\frac{\partial \Pi}{\partial t}) + \frac{1}{2}(\frac{\partial^2 \Pi}{\partial r^2})\tau^2\}dt \tag{3.8}$$

Thus our portfolio Π is now risk-less, in the sense that between t and $t+dt$ there is no random disturbance that affects the value of this portfolio, as we hedge it away [7]. So it follows from the no arbitrage principle that our portfolio Π must now earn the risk free rate r_t between t and $t + dt$. We now have that

$$\frac{d\Pi_t}{\Pi_t} = r_t dt \qquad \text{(No arbitrage)} \tag{3.9}$$

The above no arbitrage condition supplies us with an equation whose solution is the solution of the bond price. To achieve this, let's multiply all through equation (3.9) by Π and substitute the right-hand side of equation (3.8) for $d\Pi$ to obtain

$$\frac{\partial \Pi}{\partial t} + \frac{1}{2}\frac{\partial^2 \Pi}{\partial r^2}\tau^2 = r_t \Pi_t \tag{3.10}$$

From equation (3.3), we have that

$$\frac{\partial \Pi}{\partial t} = \frac{\partial P_1}{\partial t} - \triangle \frac{\partial P_2}{\partial t} \tag{3.11}$$

$$\frac{\partial^2 \Pi}{\partial r^2} = \frac{\partial^2 P_1}{\partial r^2} - \triangle \frac{\partial^2 P_2}{\partial r^2} \tag{3.12}$$

By substituting equations (3.11) and (3.12) in equation (3.10) we have that

$$\frac{\partial P_1}{\partial t} - \triangle \frac{\partial P_2}{\partial t} + \frac{1}{2}(\frac{\partial^2 P_1}{\partial r^2} - \triangle \frac{\partial^2 P_2}{\partial r^2})\tau^2 = r_t \tag{3.13}$$

$$(\frac{\partial P_1}{\partial t} + \frac{1}{2}\frac{\partial^2 P_1}{\partial r^2}\tau^2 - r_t P_1) = \triangle(\frac{\partial P_2}{\partial t} + \frac{1}{2}\frac{\partial^2 P_2}{\partial r^2}\tau^2 - r_t P_2) \tag{3.14}$$

Since

$$\triangle = \frac{\partial P_1/\partial r}{\partial P_2/\partial r}$$

it finally follows that

$$\frac{(\frac{\partial P_1}{\partial t} + \frac{1}{2}\frac{\partial^2 P_1}{\partial r^2}\tau^2 - r_t P_1)}{\partial P_1/\partial r} = \frac{\triangle(\frac{\partial P_2}{\partial t} + \frac{1}{2}\frac{\partial^2 P_2}{\partial r^2}\tau^2 - r_t P_2)}{\partial P_2/\partial r} \tag{3.15}$$

The above equation is the principal connection across interest rate securities that originates from the no arbitrage condition above. We observe that the formulas on both sides of equation (3.15) are equivalent, but applied to two different bonds. The numerator of each of the expressions in equation (3.15) gives the annualized capital return due to the passage of time and the denominator provides the "risk" of the long position in the security, expressed in terms of its sensitivity to changes in interest rates [7].

The Fundamental Pricing Equation

For every interest rate security with price $P(r,t;T)$, we have that

$$\frac{(\frac{\partial P_1}{\partial t} + \frac{1}{2}\frac{\partial^2 P_1}{\partial r^2}\tau^2 - r_t P_1)}{\partial P_1/\partial r} = -c^*(r,t) \tag{3.16}$$

where $c^*(r,t)$ is the name of the common value which is common across all the interest rate securities since all interest rate securities must satisfy equation (3.15). The negative sign in front of $c^*(r,t)$ is for convenience only, so that we can rewrite this equation as the following stochastic differential equation (S.D.E)

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial r}c^*(r,t) + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\tau^2 = rP. \tag{3.17}$$

where

$$c^*(r,t) = \eta(\xi - r(t)) - \lambda\eta_r \tag{3.18}$$

and the λ is a constant risk premium [6]. The bond price $P(r,t;T)$ is then the solution of this S.D.E. That is, it is a function of time t and interest rate r such that

(i) if we take its partial derivatives and substitute them into the left-hand side of equation (3.17), we obtain the right-hand side and,

(ii) it satisfies the boundary condition

$$P(r,T;T) = 1 \tag{3.19}$$

Note that in the above derivation, nothing restrains $P(r,t;T)$ to be a zero coupon bond. Specifically, $P(r,t;T)$ could be the price of an option, a portfolio of securities, or an option. Any of this interest rate securities must satisfy equation (3.16), and thus the S.D.E (3.17) [7]. What distinguishes one interest rate security from another is the final condition at maturity T. For example, if $P(r,t;T)$ is a zero coupon bond, the final condition is as given in equation (3.19), but if $P(r,t;T)$is the value at time t of an interest rate option with maturity T and the strike rate S_k, for example, then its price must also satisfy the S.D.E (3.17) but the final condition would now be given as

$$P(r,T;T) = N \times \max(r - S_k, 0) \tag{3.20}$$

Equation (3.17) is thus called *fundamental pricing equation* because all of the interest rate securities must satisfy it and it is important to acknowledge that this fundamental pricing equation depends on the choice of the model. A different model will yield a slightly different equation [7].

Available are a good number of term structure models that practitioners make use of, and in order to understand these pricing models, it is important to have a clue of how they are used. Term structure models are in general of interest to bond traders, to plain-vanilla-option traders and to complex-derivatives traders. It is important to bond traders because if the yield curve dynamics has description that were correct, then failure of the model to match the market bond prices exactly could speculate trading opportunities. In a nutshell, term structure models have be used in various ways by various market practitioners (traders) of which complex practitioners have become the largest group. As a result, few practitioners rely on equilibrium or arbitrage-free models now-a-days to estimate the relative value of the bond being traded, and moreover, much attention has been drawn towards the pricing of complex productions, which requires an accurate description of the underlying inputs (that is, bond prices and yield volatilities) [26,34]. In general, a trader want a model which is

- flexible enough to cover most situations arising in practice;
- simple enough that one can compute answers in reasonable time;
- well-specified, in that required inputs can be observed or estimated;
- realistic, in that the model will not do silly things;
- a good fit of the model to data;
- an equilibrium derivation of the model.

In the following section, we present the various categories these model can belong to.

3.1 Model Taxonomy.

To have a good understanding of these models, it is important to point out a number of the distinctive characteristics that dwell among them. These are of great relevance to practitioners wishing to apply valuation algorithms.

3.1.1 Continuous versus Discrete Models.

One very important question that arise is whether to model the term structure dynamics in a discrete or in a continuous framework.

Most of the interest rates models available in the literature were stated in a continuous time framework. The strength of continuous time stochastic calculus enables the derivations, proofs, and provides an adequate framework to produce more precise theoretical solutions and more improved empirical hypotheses for these models.

On the space dimension, diffusion models were in lead until recently. But should we use a diffusion model, or allow for discontinuity? Models based on jumps or point processes have appeared very recently in order to model discontinuous real world phenomena such as the central bank intervention [6].

3.1.2 Bond prices, Interest Rate versus Yield Curve Models.

Even after identifying whether to consider a discrete or a continuous framework, the question of what to model now arises.

The early models of the term structure tried to model the dynamics of the bond price. However, results obtained with this approach this not give room for better understanding of the term structure, which is hidden in the bond price [6].

Most of the interest rate models are simple models of the stochastic growth of a given interest rate that is often chosen to be the short rate. This interest rate is often defined as Markovian that is, its future growth only depends on its current value and not on the historical path it followed to get there. This translates the valuation problem into a partial differential equation that can be solved analytically or numerically.

An alternative approach is to specify the stochastic behavior of the entire term structure of interest rates, either by using all yields or all forward rates. Attractive as this approach is, never-the-less increases the complexity of the model and as a consequence, this has hindered whole yield curve models from becoming more popular.

3.1.3 Single versus Multi-Factor Models.

The term structure of interest rates is driven by a set of factors (variables) in factor models. Most empirical studies using the principal component analysis (PCA) have decomposed the movement of the term structure of interest rate into three independent and non-correlated factors [7].

The first and the largest of these principal components (factor) is the *shift* of the term structure. This is a parallel movement of all the rates and it accounts for up to 80%-90% of the total variance. As a result, valuation can be reduced to a one factor in many instances with little loss of accuracy. Consequently, only securities whose payoffs are primarily sensitive to the shape of the spot curve rather than its overall level will not be modeled well with this approach.

The second of these principal components (factor) is the *twist* of the term structure. This is a situation in which the long rates and the short term rates move in opposite directions. This usually accounts for 5%-10% of the total variance.

The third principal component (factor) is called the *butterfly*. In this case, the immediate rate moves in the opposite direction of the short and the long term rate and it usually accounts for 1%-2% of the total variance.

It is possible to move from a one factor model to a multi-factor model although the practical implementation of these multi-factors models can be complicated and slow and as such, there may be some benefits of instead using a one-factor model when possible. However, for the single factor models, it should be pointed out that it does not necessarily imply that the whole term structure is forced to move in parallel, but instead that the one single source of uncertainty is sufficient to explain the movements of the terms structure. For example, almost all arbitrage-based single factor models assume the instantaneous spot rate to be the single state variable whereas some securities are sensible to the shape of the term structure and not only to its level and as such they will require at least a two factor model [7].

3.1.4 Fitted versus Non Fitted Models.

In the case of fitted models, the term structure is determined exogenously, generally using market data, and the SDE of some state variables is specified such that this term structure is obtained at a particular date. Usually, the term structure here is that of interest rates, forward rates, yield volatilities etc. Ideally, these models are built to particularly match an arbitrary exogenous initial term structure.

On the contrary, the dynamics of the state variables of the non fitted models have to be specified first. And so based on a particular specification, a given term structure is obtained endogenously. As a result, these models do not in general match well the initially observed term structure.

3.1.5 Endogenous versus Exogenous Models.

The very former rate models were not built so as to match an arbitrary initial term structure. Instead, because of the quest for analytical simplicity, models such as the Vasicek and the Cox-Ingersoll-Ross (CIR) models accommodate some constant parameters that define an endogenously prescribed term structure. That is, the initial spot curve is given by an analytical formula in terms of the model parameters. Because these models hypothesis yield curves derived from an assumption of economy equilibrium based on a given market price of risk and other parameters governing collective expectations, they are some times termed *"equilibrium"* models. These models have too few degrees of freedom to truly represent real markets and as a result, they are not particularly useful as the basis for valuation algorithms. To be used for valuation, a model (or these equilibrium models) needs to be calibrated to the initial spot rate curve. That is, the model structure must harbor an exogenously determined spot rate curve which is customary given by fitting to bond prices, or sometimes to future prices and swap rate. All models in common use are of this type [24].

3.1.6 Arbitrage Free versus Equilibrium Models.

Arbitrage-Free models begin by setting up assumptions about the stochastic behavior of one or many interest rates and about a specific market price of risk and then derive the price of all contingent claims assuming that there are no arbitrage opportunities on the market (that is, there is no risk free financial strategy with zero setup cost that should give with certainty a positive return).

On the contrary, equilibrium models start from the description of the economy to then derive the term structure of interest rates, the risk premium, and other assets prices endogenously, assuming that the market is at equilibrium.

Never-the-less, the distinction is delicate, since equilibrium models should be arbitrage free (otherwise, the economy would not be at equilibrium), and as some so-called *"arbitrage-free"* models were shown later on to allow for arbitrage opportunities [6].

Chapter FOUR

Single Factor Equilibrium Interest Rate Models

These models are called *single factor models* due to the fact that they present a single source of randomness. The early equilibrium models are based on a mathematical model of the economy. They focus on describing and explaining the interest rate term structure. Equilibrium models are also referred to as *"endogenous term structure models"* because the term structure of interest rate is an output of, rather than an input to these models. If we have the initial zero-coupon bond curve from the market, the parameters of the equilibrium models are chosen such that the models produce a zero-coupon bond as closes as possible to that of the one observed in the market. However, a fundamental problem with this equilibrium approach is that the models cannot exactly reproduce the initial yield curve. As a consequence, most practitioners have very little confidence in using these models. Below are some well renowned one factor equilibrium interest rate models.

4.1 Vasicek Mean Reversion Model.

Vasicek in 1977 [1,2] introduced a new theory into the financial econometric world by capturing the mean reversion property of interest (that is, the expected value of the short rate tends to a constant value ξ with velocity depending on η as time grows, while its variance does not explode) rate while relating the movement to previous time. This mean reversion property is one of the key innovations of the model and an economic justification of this is that;

High interest rates tend to cause the economy to slow down as borrowers require less funds. This causes the rates to decline to the equilibrium long-term mean. However, when rates are low, funds are of high demand on the part of the borrowers making the rates to tend to increase again towards the long-term mean.

Vasicek derived the general form of the term structure of interest rates to model the evolution of the shor-term interest rate. His work is the first account of a bond pricing model that incorporates stochastic interest rate in which he analyses pure (zero-coupon) discount bond. This was one of the first model to make a significant impact on interest rate modeling. Even though his paper was titled *"An Equilibrium Characterization of the Term Structure"* he did not make use of any assumptions about equilibrium within an elementary economy, nor did he make use of an equilibrium argument in the derivation. Rather, his derivation relied on an arbitrage argument, similar to the one employed by Black and Scholes in their option pricing model. Vasicek instead made assumptions about the stochastic evolution of interest rates by exogenously specifying the process describing the short-term interest rate.

Vasicek assumes that the instantaneous interest rate follows an Ornstein-Urlenbeck process(also known as the mean reverting process):

$$dr(t) = \eta(\xi - r)dt + \tau d\Phi \tag{4.1}$$

$$r(0) = r_0 \tag{4.2}$$

The deterministic part of the process which is the instantaneous drift $\eta(\xi - r)$ denotes the force that keeps on dragging the short-rate towards its long-term mean ξ with a speed η proportional to the deviation of the process from the mean. $\tau d\Phi$ is the stochastic part having a constant instantaneous variance τ^2 and causes the process to fluctuate around the level ξ in a haphazard but continuous way. $d\Phi$ is itself a standard Wiener process, that is, $d\Phi$ is normally distributed with mean 0 and standard deviation $\sqrt{dt}$.

The stochastic differential equation (SDE) (4.1) is a linear equation with constant volatility. This implies that the short-rate is both Gaussian and Markovian (that is, the system has no memory. Future developments of the short-rate are independent of the past movements). In particular, it is a continuous Markovian process and therefore a diffusion process.

Being a Markov process with normally distributed increments, and in contrast to the random walk (Wiener process) which is unstable and diverges to infinite values after long time, the Ornstein-Urlenbeck process possesses a stationary distribution [3].

Based on the current value $P(t, T)$ (price of the default-free pure discount bond) where t is the current time or period, T is the bond's maturity date and $T - t$ is the term to maturity, Vasicek made the following three assumptions in his model:

(1) The current sport interest rate is known with certainty but however, subsequent values of the spot rate are not known. $r(t)$ follows a stochastic process which is a continuous function of time and follows a Markov process. This implies that the spot rate process is fully characterized by a single state variable.

(2) The development of the spot rate over (t, T) is fully determined by its current value $r(t)$ and so the bond price may be written as a function of the current spot rate:

$$P(t, T) = P(r(t), t, T)$$

Hence, the entire term structure is determined by the spot rate.

(3) The market is assumed to be efficient. This implies

   (i) there are no transaction costs,

   (ii) information is simultaneously distributed to all investors,

   (iii) investors are rational with homogeneous expectations,

   (iv) profitable, risk-less arbitrage is not possible.

A special feature of this model is that it has a closed form solution and in order to get this solution, we are going to employ the following approach, that is, a standard calculus base approach.

Consider the homogeneous differential equation when $\xi = \tau = 0$. Then we can write equation (4.1) as follows;

$$
\begin{aligned}
dr(t) &= \eta r(t)dt \qquad (4.3)\\
\implies \int \frac{dr(t)}{r(t)} &= -\int \eta dt\\
\implies \ln|r(t)| &= -\eta t + B\\
\implies |r(t)| &= e^{B}e^{-\eta t}\\
\implies r(t) &= \pm e^{-\eta t}e^{B}\\
\implies r(t) &= Ke^{-\eta t} \qquad (4.4)
\end{aligned}
$$

We now see a particular solution of the non-homogeneous equation in the form (4.3) with K replaced by an unknown function μ(t),

$$
\begin{aligned}
r_1(t) &= \mu(t)e^{-\eta t} \qquad (4.5)\\
\implies \mu(t) &= r_1(t)e^{\eta t} \qquad (4.6)
\end{aligned}
$$

By taking the derivative on both sides of equation (4.6), we have the following;

$$
\begin{aligned}
d\mu(t) &= dr_1(t)e^{\eta t}\\
&= dr_1(t)e^{\eta t} + \eta e^{\eta t}r_1(t)dt\\
&= \eta(\xi - r)e^{\eta t}dt + \tau e^{\eta t}d\Phi + \eta e^{\eta t}r_1(t)dt\\
&= \eta\xi e^{\eta t}dt - \eta r_1(t)e^{\eta t}dt + \tau e^{\eta t}d\Phi + \eta e^{\eta t}r_1(t)dt\\
&= \eta\xi e^{\eta t}dt + \tau e^{\eta t}d\Phi
\end{aligned}
$$

Thus μ(t) satisfies the ordinary differential equation (O.D.E)

$$d\mu(t) = \eta\xi e^{\eta t}dt + \tau e^{\eta t}d\Phi(t) \qquad (4.7)$$

thus

$$
\begin{aligned}
\int d\mu(t) &= \int \eta\xi e^{\eta t}dt + \int \tau e^{\eta t}d\Phi(t)\\
\implies \mu(t) &= \xi e^{\eta t} + \tau \int_0^t e^{\eta s}d\Phi(s) \qquad (4.8)
\end{aligned}
$$

Hence, the solution to our problem is the sum of equation (4.4) and (4.8). That is

$$r(t) = Ke^{-\eta t} + \xi e^{\eta t} + \tau \int_0^t e^{\eta s}d\Phi(s) \qquad (4.9)$$

In an attempt to impose the initial condition (4.2), we Set $K = r_0 - \xi$ such that

$$
\begin{aligned}
r(t) &= r_0 e^{-\eta t} + \xi - \xi e^{-\eta t} + \tau \int_0^t e^{-\eta(t-s)} d\Phi(s) \\
&= r_0 e^{-\eta t} + \xi(1 - e^{-\eta t}) + \tau \int_0^t e^{-\eta(t-s)} d\Phi(s) \\
&= e^{-\eta t}[r_0 + \xi(e^{-\eta t} - 1) + \tau \int_0^t e^{\eta s} d\Phi(s)] \\
&= \mu_t + \tau \int_0^t e^{-\eta(t-s)} d\Phi(s) \qquad (4.10)
\end{aligned}
$$

Equation (4.10) is the closed form solution to our S.D.E (4.1) where ξ is some kind of level r is trying to attain. We call this the mean-reverting level [1].

The fist part of this solution (μ_t) is deterministic and just a number while the second part is a stoch (stochastic). $r(t)$ is Gaussian with mean and variance [4,5,6]:

$$
\begin{aligned}
E(r(t)) &= E(r_0 e^{-\eta t} + \xi(1 - e^{-\eta t}) + \tau \int_0^t e^{-\eta(t-s)} d\Phi(s)) \\
E(r(t)) &= E(r_0 e^{-\eta t}) + E(\xi(1 - e^{-\eta t}) + E(\tau \int_0^t e^{-\eta(t-s)} d\Phi(s))) \qquad (4.11)
\end{aligned}
$$

Since $d\Phi(s)$ is normally distributed with mean 0 and variance 1, it follows that

$$
\implies E(r(t)) = r_0 e^{-\eta t} + \xi(1 - e^{-\eta t}) \qquad (4.12)
$$

and

$$
\begin{aligned}
V(r(t)) &= E[(\tau \int_0^t e^{-\eta(t-s)} d\Phi(s))^2] = E[(\tau e^{-\eta t} \int_0^t e^{\eta s} d\Phi(s))^2] \\
\implies V(r(t)) &= \tau^2 e^{-2\eta t} E[\int_0^t e^{2\eta s} d\Phi(s)] \\
\implies V(r(t)) &= \frac{\tau^2}{2\eta}(1 - e^{-2\eta t}) \qquad (4.13)
\end{aligned}
$$

As t$\to \infty$, we have

$$
\lim_{t\to\infty} E(r(t)) = r_0 \lim_{t\to\infty} e^{-\eta t} + \xi(1 - \lim_{t\to\infty} e^{-\eta t}) = \xi \qquad (4.14)
$$

and

$$
\lim_{t\to\infty} V(r(t)) = \frac{\tau^2}{2\eta}(1 - \lim_{t\to\infty} e^{-2\eta t}) = \frac{\tau^2}{2\eta} \qquad (4.15)
$$

We see that the Vasicek model gives rise to Gaussian mean-reverting interest rates with long term mean equal to ξ and long term variance equal $\frac{\tau^2}{2\eta}$. So $N[\xi, \frac{\tau^2}{2\eta}]$ is the limiting distribution of $r(t)$.

The order of magnitude at which random fluctuation interfering with the mean reversion takes place is $\frac{\tau^2}{\sqrt{2\eta}}$ and

$$
\frac{\tau^2}{\sqrt{2\eta}} \to 0 \qquad as \qquad \eta \to \infty
$$

and thus strongly mean reverting processes are characterized by low volatility.

Simple and straightforward as it is, we must however acknowledge that the Vasicek model has a number of genuine flaws [3]:

(1) The dependence on a single factor greatly limits the possible shapes of the yield curve and often leads into situations where the theoretical yield curve does not corresponds to the market yield curve. In fact this is the case nearly always.

(2) The most problematic point in the Vasicek model is that it can produce negative interest rates with non-zero probability.

(3) It is impossible to fit the entire forward curve as the initial condition.

(4) There is one volatility parameter only available for calibration(two, if we count the mean reversion rate). That makes fitting the volatility structure virtually impossible.

These problems have been overcome by the short rate model due to Cox, Ingersoll, and Ross [17]. Some can also be overcome by means of a slight extension of the model, that is, for example the one factor Hull-White mode [5].

4.1.1 The Vasicek Bond Pricing Formula

We now provide the bond price formula for the Vasicek model. Using the discounted expected terminal value of the bond with respect to the probability measure $\mathcal{P}$ and the filtration $\mathcal{F}_t$ [6].

$$P(t,T) = E[e^{-\int_t^T r(s)ds}|\mathcal{F}_t] \tag{4.16}$$

the solution is given to be

$$P(t,T) = e^{a(t,T)r(t)+b(t,T)} \tag{4.17}$$

where

$$a(t,T) = \frac{1}{\eta}(e^{-(T-t)\eta} - 1) \tag{4.18}$$

$$b(t,T) = \frac{\tau^2}{4\eta^3}(1 - e^{-2(T-t)\eta}) + \frac{1}{\eta}\left(\xi - \frac{\lambda\tau}{\eta} - \frac{\tau^2}{\eta^2}\right)(1 - e^{-(T-t)\eta}) - \left(\xi - \frac{\lambda\tau}{\eta} - \frac{\tau^2}{\eta^2}\right)(T-t) \tag{4.19}$$

and the bond price dynamics is given by

$$\frac{dP}{P} = \left[r(t) + \frac{\lambda\tau}{\eta}(e^{-(T-t)\eta} - 1)\right]dt + \frac{\lambda\tau}{\eta}(e^{-(T-t)\eta} - 1)d\Phi(t) \tag{4.20}$$

which implies that bond prices are log-normally distributed. It follows that the price of a coupon bond with maturity T and coupon rate c is given by [7]

$$P_c^*(r,t;T) = \frac{100 \times c}{2}\sum_{i=1}^{n} P(r,t;T_i) + 100 \times P(r,t;T_n) \tag{4.21}$$

Given a discount factor function $P(r,t;T)$, as a function of T, we can then compute the zero coupon spot rate , that is, the term structure of interest rates. For this purpose, lets clarify the following concepts: The bond pricing formula in equation (4.17) is written as a function of interest rate r, the current rate t, and the maturity date T. However, it hides the fact that the value of the zero coupon bond at time t depends only on *time to maturity* $\tau^* = T-t$. Note that when we want to emphasize the time to maturity aspect of a bond, we will use the following notation

$$P(r_t,\tau^*) = P((r_t,t;T) \tag{4.22}$$

$$a(\tau^*) = a(0;T-t) \tag{4.23}$$

$$b(\tau^*) = b(0;T-t) \tag{4.24}$$

4.1.2 The Term Structure of Interest Rates Under the Vasicek Model

We can denote the term structure of interest rates as a function of the current interest rate $r_t(\tau^*)$ and time to maturity τ^*. This is given by [6,7]

$$\begin{aligned} r_t(\tau^*) &= -\frac{\ln(P(r_t,\tau^*))}{\tau^*} \\ &= -\frac{a(\tau^*)}{\tau^*} + \frac{b(\tau^*)}{\tau^*} r_t \end{aligned} \tag{4.25}$$

$$\begin{aligned} r_t(\tau^*) = -\frac{1}{\tau^*}\Bigg[&\frac{1}{\eta}(e^{-\tau^*\eta}-1)r(t) + \frac{\tau^2}{4\eta^3}(1-e^{-2\tau^*\eta}) \\ &+ \frac{1}{\eta}\left(\xi - \frac{\lambda\tau}{\eta} - \frac{\tau^2}{\eta^2}\right)(1-e^{-\tau^*\eta}) - \left(\xi - \frac{\lambda\tau}{\eta} - \frac{\tau^2}{2\eta^2}\right)\tau^*\Bigg] \end{aligned} \tag{4.26}$$

That is, each point on the term structure is a linear function of the current short-term interest rate r_t. As r_t moves over time according to the Vasicek model, so does the whole term structure of interest rates. Figure [4.1] plots three spot curves for estimated parameters. The figure shows that if the current short-term interest rate r_t is low, the Vasicek model implies a rising term structure of the interest rates. If instead the current short-term interest rate is high, the Vasicek model implies a decreasing term structure. Note that as the interest rate moves up and down, the term structure moves too. Especially, the *level*, *slope* and *movement* change over time [7].

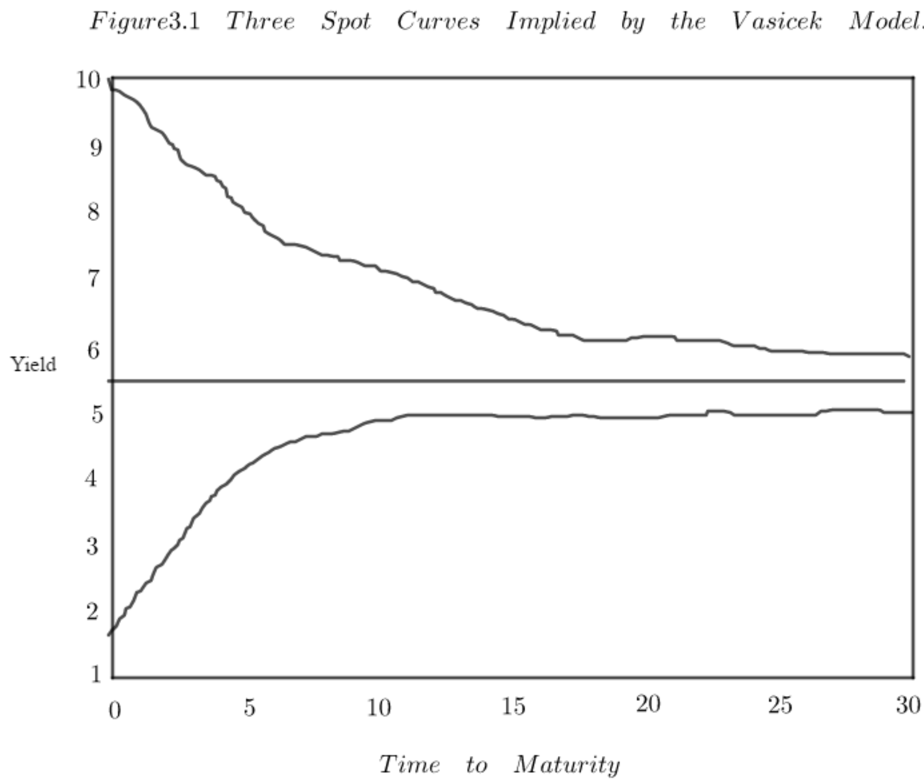


Figure 4.1: Three spot curves implied by the Vasicek model

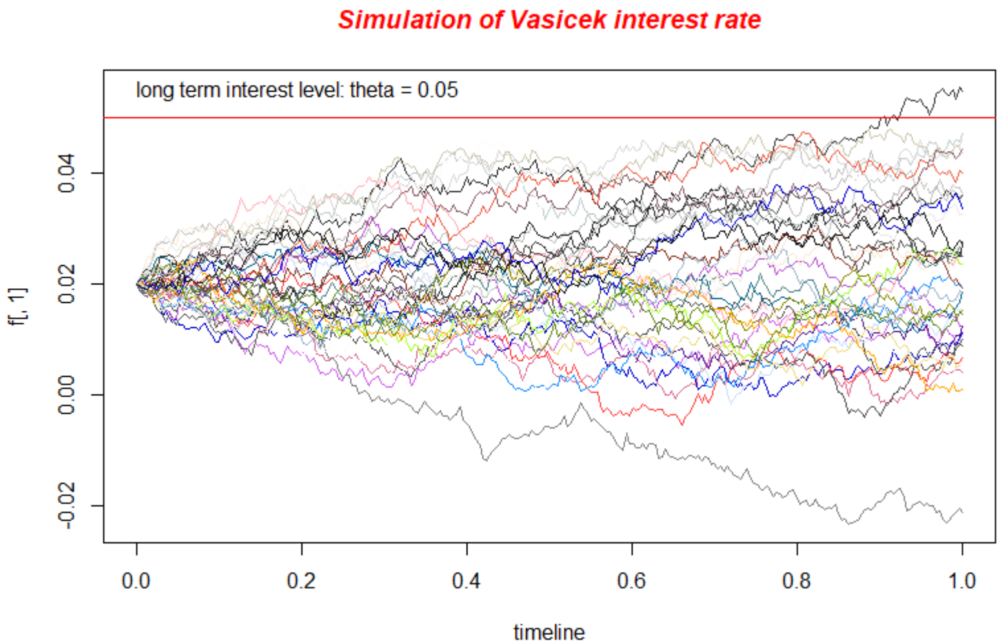


Figure 4.2: Graph of the Vasicek short rate $t \to r(t)$

As can be seen from the figure above, the value of $r(t)$ in the Vasicek model may become negative due to its Gaussian distribution. This is a major disadvantage of the model even though real interest rate may fall below zero.

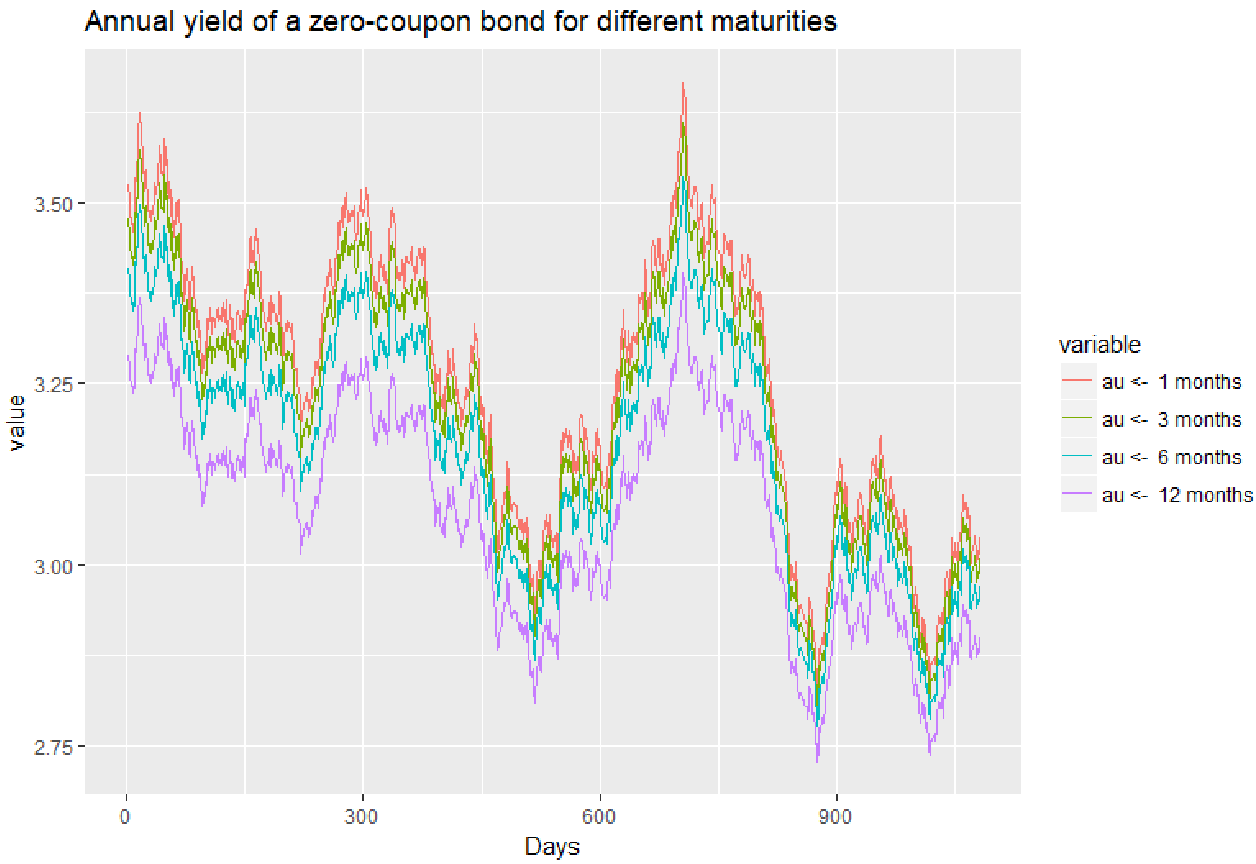


Figure 4.3: (b)

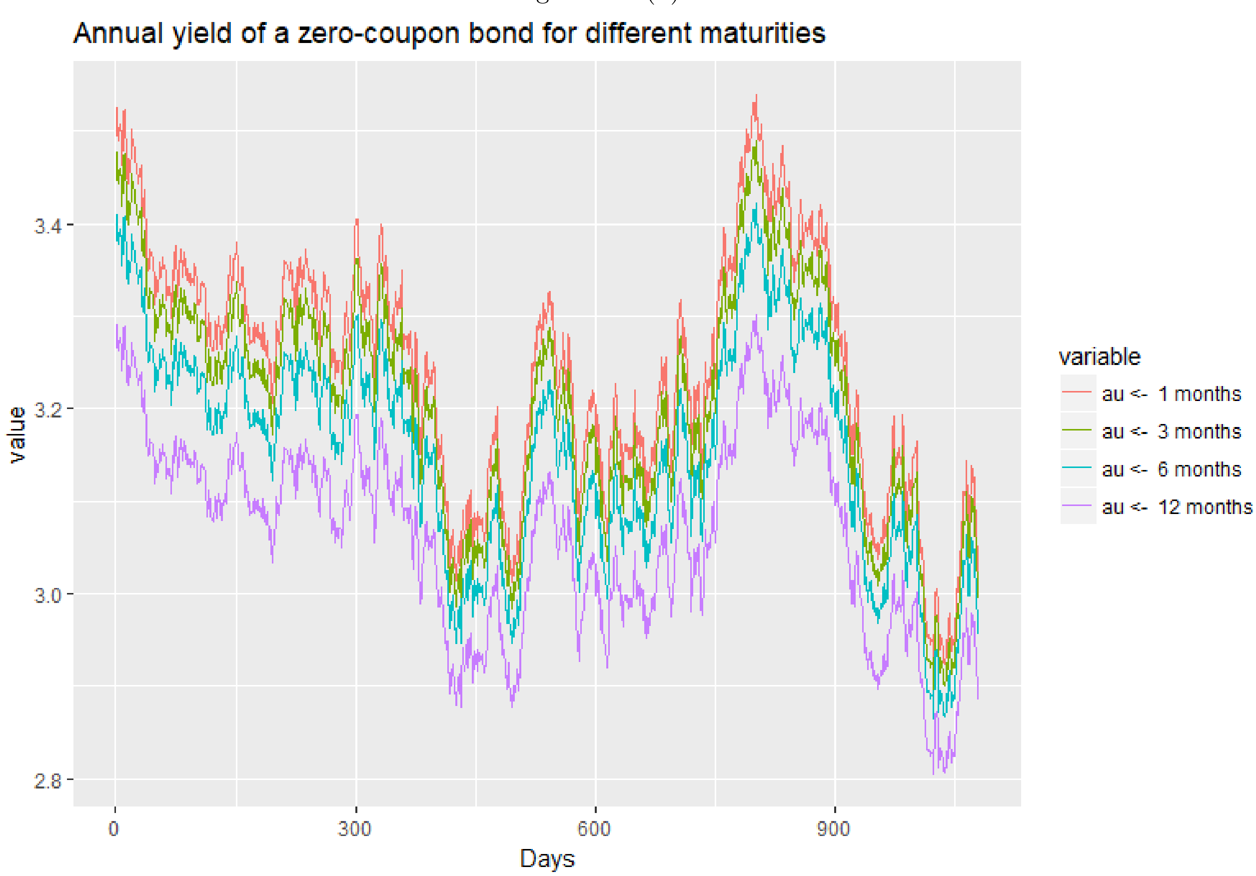


Figure 4.4: (a)

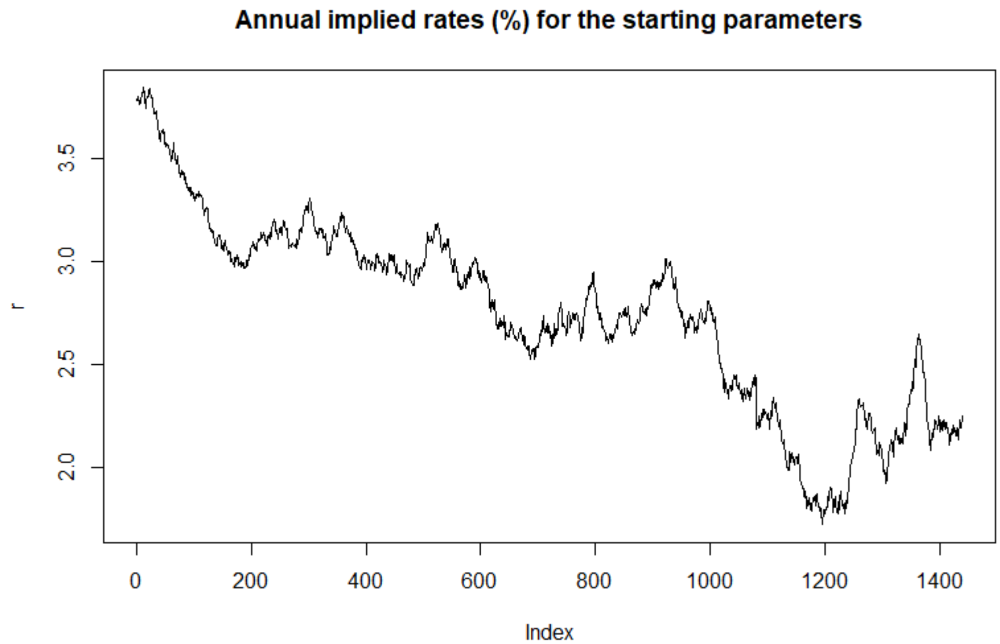


Figure 4.5: (a)

4.2 Cox, Ingersoll, Ross Model.

One of the most important term structure model in the literature is the **John C. Cox, Jonathan E. Ingersoll and Stephen A. Ross (CIR)** introduced in 1985 [9]. Contrary to Vasicek, their approach began with a thorough stipulation of an equilibrium economy. Assumptions were made about the stochastic evolution of exogenous state variables and about investors preferences. The form of the short term interest rate process and the prices of contingent claims were then endogenously derived from within the equilibrium economy. Since the bond prices are derived from exogenous specifications of the economy (that is, production opportunities, investors' tastes and beliefs about future states of the world), the CIR model is considered a complete equilibrium model.

This general equilibrium endogenous approach to the short rate process let to the introduction of a *"square-root"* term in the diffusion coefficient of the instantaneous short rate dynamics, thus, the model also goes under the name of the *square root model* of interest rates [7].

The model provides solutions for bond prices and a complete characterization of the term structure which incorporates risk premiums and expectations for future interest rates. It has been an alternative to the Vasicek model and a benchmark for many years because of the fact that, contrary to the Vasicek model, the instantaneous short rate is always positive. For the purpose of term structure modeling, CIR imposed the following assumptions about their specialized equilibrium economy [8]:

(1) The change in production opportunities over time is determined by a single state variable, say M.

(2) The mean and variance of the rate of return on the production processes are proportional to M. Hence, the state variable M determines the rate of evolution of capital and neither the mean nor variance will dominate the portfolio decisions for increasing value of M.

(3) The state variable X follows the following stochastic process:

$$dX(t) = [kX + b]dt + c\sqrt{X}d\Phi(t) \tag{4.27}$$

are K and b are constants, $b \geq 0$ and c a vector of constants.

By taking in the above assumptions into their economic model, CIR [8] derived an explicit formula for the equilibrium interest rate in terms of the state variable X, the means and variances of the rates of return on the production processes in the economy and the parameters of its stochastic process. Computing the drift and variance of this equilibrium interest rate and by designating a new Wiener process $\Phi(t)$, such that:

$$\tau\sqrt{r}d\Phi \equiv c\sqrt{X}dW(t) \tag{4.28}$$

CIR describe the dynamics of the short interest rate, r as government by the following S.D.E

$$dr(t) = \eta(\xi - r)dt + \tau\sqrt{r}d\Phi(t) \tag{4.29}$$

$$r(0) = r_0 \geq 0 \tag{4.30}$$

where Φ is a standard Wiener process, η, ξ, τ are positive constants and the risk premium at equilibrium is shown to be

$$\lambda(r, t) = \lambda\sqrt{r(t)} \tag{4.31}$$

Equation (4.29) symbolizes a continuous time first order autoregressive process where the stochastic interest rate is dragged to its long term mean ξ with speed η (mean reverting, that is, Ornstein-Uhlenbeck process) and enforcing the additional constraint that $2\xi\eta \geq \tau^2$ ensures that the interest rate cannot become negative.

The model is slightly less tractable than the Vasicek model, in the sense that analytical formulas for some securities are not available. However, since interest rates are always positive in this model, this has been considered an important step forward in the term structure modeling. The reason why interest rates are always positive in this model can be seen from observing the diffusion term $\tau\sqrt{r}$ in equation (4.29). When the interest rate r is moving towards zero, the diffusion part $\tau\sqrt{r}$ declines, and it becomes in fact zero when r hits zero. When $r = 0$, the only term left in equation (4.29) is

$$dr(t) = \eta\xi > 0 \tag{4.32}$$

The interest rate behavior implied by this structure thus has the following empirically relevant properties:

(1) negative interest rates are excluded,

(2) if the interest rate reaches zero, it can subsequently become positive,

(3) the absolute variance of the interest rate increases when the interst rate itself increases.

(4) there is a steady state distribution of the interest rate.

It turns out that the distribution of {r} is non-central Chi-square with first and second moments as illustrated below.
To compute the first and second moments of the rate of interest under the CIR model, lets consider the integral form of equation (4.29)

$$r(t) = r_0 + \eta \int_0^t (\xi - r(s))ds + \tau \int_0^t \sqrt{r(s)} d\Phi(s) \tag{4.33}$$

since the initial rate of interest r_0 is known. By taking expectations of the above equation and coupled to the fact that a Wiener process has zero mean, we have

$$E[r(t)|r_0] = r_0 + \eta \int_0^t (\xi - E(r(s)|r_0))ds \tag{4.34}$$

Representing this equation in differential form leads to

$$\frac{\partial}{\partial t} E[r(t)|r_0] = \eta(\xi - E(r(s)|r_0)) \tag{4.35}$$

Equation (4.35) is an ordinary differential equation which may be solved by the method of separation of variables as follows

$$\frac{dE[r(t)|r_0]}{\xi - E(r(s)|r_0)} = \eta ds \tag{4.36}$$

$$\Longrightarrow \int_0^t \frac{dE[r(t)|r_0]}{\xi - E(r(s)|r_0)} = \int_0^t \eta ds \tag{4.37}$$

$$\Longrightarrow \ln\left(\frac{\xi - E(r(s)|r_0)}{\xi - r_0}\right) = -\eta(t - 0) \tag{4.38}$$

$$\Longrightarrow E(r(s)|r_0) = \xi + (r_0 - \xi)e^{-\eta t} \tag{4.39}$$

which is indeed the first moment of the interest rate under the CIR model. Also,

$$\begin{aligned}
\lim_{t\to\infty} E(r(s)|r_0) &= \lim_{t\to\infty} \left[\xi + (r_0 - \xi)e^{-\eta t}\right] \\
\lim_{t\to\infty} E(r(s)|r_0) &= \xi + (r_0 - \xi) \lim_{t\to\infty} e^{-\eta t} = \xi
\end{aligned} \tag{4.40}$$

In the case of the second moment, lets define a function $f(y) = y^2$, then by applying Ito's formula to calculate the the stochastic differential equation satisfied by $r^2(t)$, we have that

$$\begin{aligned}
df(r(t)) &= f^{'}(r(t))dr(t) + \frac{1}{2} f^{''}(r(t))dr(t) \\
d(r^2(t)) &= 2r(t)\left[\eta(\xi - r(t))dt + \tau\sqrt{r} d\Phi\right] + \left[\eta(\xi - r(t))dt + \tau\sqrt{r} d\Phi\right]^2 \\
d(r^2(t)) &= 2\eta\xi r(t)dt - 2\eta r^2(t)dt + 2\tau r^{\frac{3}{2}}(t)d\Phi + \tau^2 r(t)dt \\
d(r^2(t)) &= (2\eta\xi + \tau^2)r(t)dt - 2\eta r^2(t)dt + 2\tau r^{\frac{3}{2}}(t)d\Phi
\end{aligned}$$

so

$$r^2(t) = r_0^2 + (2\eta\xi + \tau^2)\int_0^t r(s)ds - 2\eta\int_0^t r^2(s)ds + 2\tau\int_0^t r^{\frac{3}{2}}(s)d\Phi(s)$$

Taking the conditional expectation of $r^2(t)$ we get

$$E[r^2(t)|r_0] = r_0^2 + (2\eta\xi + \tau^2)\int_0^t E[r(s)|r_0]ds - 2\eta\int_0^t E[r^2(s)|r_0]ds$$

By now taking the partial differentiation with respect to t yields

$$\frac{\partial}{\partial t}E[r^2(t)|r_0] = (2\eta\xi + \tau^2)E[r(s)|r_0] - 2\eta E[r^2(s)|r_0]$$

and hence

$$\begin{aligned}
\frac{\partial}{\partial t}\big(e^{2\eta(t-0)}E[r^2(t)|r_0]\big) &= 2\eta e^{2\eta(t-0)}E[r^2(s)|r_0] + e^{2\eta(t-0)}\frac{\partial}{\partial t}E[r^2(t)|r_0] \\
&= 2\eta e^{2\eta(t-0)}E[r^2(t)|r_0] + 2\eta\xi e^{2\eta(t-0)}E[r(t)|r_0] \\
&+ \tau^2 e^{2\eta(t-0)}E[r(t)|r_0] - 2\eta e^{2\eta(t-0)}E[r(t)|r_0] \\
&= (2\eta\xi + \tau^2)e^{2\eta(t-0)}E[r(t)|r_0] \qquad (4.41)
\end{aligned}$$

Integrating equation (4.41) over the closed interval $[0, t]$ and using equation (4.39) above gives

$$\begin{aligned}
e^{2\eta(t-0)}E[r^2(t)|r_0] - r_0^2 &= \int_0^t (2\eta\xi + \tau^2)e^{2\eta(s-0)}E[r(s)|r_0]ds \\
&= \int_0^t (2\eta\xi + \tau^2)e^{2\eta(s-0)}\big(r_0 e^{-\eta(s-0)} + \xi(1 - e^{-\eta(s-0)})\big)ds \\
&= (2\eta\xi + \tau^2)r_0\int_0^t e^{\eta(t-0)}ds + \xi(2\eta\xi + \tau^2)\int_0^t \big(e^{2\eta(t-0)} - e^{\eta(t-0)}\big)ds \\
&= (2\eta\xi + \tau^2)(r_0 - \xi)\int_0^t e^{\eta(s-0)}ds + \xi(2\eta\xi + \tau^2)\int_0^t e^{2\eta(s-0)}ds \\
&= \frac{1}{\eta}(2\eta\xi + \tau^2)(r_0 - \xi)(e^{2\eta(t-0)} - 1) + \frac{\xi}{2\eta}(2\eta\xi + \tau^2)(e^{2\eta(t-0)} - 1) \\
E[r^2(t)|r(t)] &= r^2(t)(e^{-2\eta(t-0)}) + \frac{1}{\eta}(2\eta\xi + \tau^2)(r_0 - \xi)e^{-\eta(t-0)} - \frac{1}{\eta}(2\eta\xi + \tau^2)(r_0 - \xi)e^{-2\eta(t-0)} \\
&+ \frac{1}{2\eta}\xi(2\eta\xi + \tau^2) - \frac{1}{2\eta}\xi(2\eta\xi + \tau^2)e^{-2\eta(t-0)} \\
&= \frac{\xi\tau^2}{2\eta} + \xi^2 + (r_0 - \xi)(\frac{\tau^2}{\eta} + 2\xi)e^{-\eta(t-0)} + (r_0^2 + \xi^2 - 2\xi r_0)e^{-2\eta(t-0)} \\
&+ (\frac{\tau^2\xi}{2\eta} - \frac{\tau^2}{\eta}r_0)e^{-2\eta(t-0)} \\
&= \frac{\xi\tau^2}{2\eta} + \xi^2 + (r_0 - \xi)(\frac{\tau^2}{\eta} + 2\xi)e^{-\eta(t-0)} + (r_0 - \xi)^2 e^{-2\eta(t-0)} + \frac{\tau^2}{\eta}\left(\frac{\xi}{2} - r_0\right)e^{-2\eta(t-0)}
\end{aligned}$$

Thus, using equation (4.39) we have

$$(E[r(t)|r(t)])^2 = (r_0 - \xi)^2 e^{-2\eta(t-0)} + \xi^2 + 2\xi(r_0 - \xi)^2 e^{-\eta(t-0)}$$

and so the conditional variance of $r(t)$ is:

$$\begin{aligned}
&Var(r(t)|r_0) = E[r^2(t)|r_0] - (E[r(t)|r(t)])^2 \\
\Longrightarrow\ &Var(r(t)|r_0) = \frac{\xi\tau^2}{2\eta} + \frac{\tau^2}{\eta}(r_0 - \xi)e^{-\eta(t-0)} + \frac{\tau^2}{\eta}(\frac{\xi}{2} - r_0)e^{-2\eta(t-0)} \qquad (4.42)
\end{aligned}$$

Also,

$$\lim_{t\to\infty} Var(r(s)|r_0) = \lim_{t\to\infty}\left[\frac{\xi\tau^2}{2\eta} + \frac{\tau^2}{\eta}(r_0-\xi)e^{-\eta t} + \frac{\tau^2}{\eta}(\frac{\xi}{2} - r_0)e^{-2\eta t}\right]$$
$$= \frac{\xi\tau^2}{2\eta} + \frac{\tau^2}{\eta}(r_0-\xi)\lim_{t\to\infty} e^{-\eta t} + \frac{\tau^2}{\eta}\big(\frac{\xi}{2} - r_0\big)\lim_{t\to\infty} e^{-2\eta t}$$
$$= \frac{\xi\tau^2}{2\eta} \tag{4.43}$$

As can be seen, while the Vasicek model gives positive probability to negative interest rates, the CIR model cuts out the negative part of the distribution. The non-central Chi-square distribution is positively skewed, giving some positive probability to high interest rates [7,17]. Although the CIR model is mainly used in finance in modeling interest rates, it should be noted that this process has other financial applications. One drawback of The S.D.E (4.29) is that it is not explicitly solvable, hence the tractability of the CIR model is not as good as the Vasicek model in this regard. In practical usage of such models (for example, to price options) we are often faced with the problem of simulating a CIR process. In general there are two ways to do it, namely, **exact simulation method** and **approximation schemes**. However, there are pros and cons associated with each method. We also note that there are other ways (model) of ensuring the short rate from being negative such as the exponential Vasicek model, Hull and White model, Black and Karasinski etc.

4.2.1 Bond Price Under the Cox, Ingersoll, and Ross Model

Similar to the Vasicek model, the CIR model also has closed form solution for the zero coupon bonds. Following the same steps used in the Vasicek model, the CIR model as any interest rate security must satisfy the fundamental pricing equation (3.17) and in particular, a bond price must satisfy this P.D.E under the boundary condition $P(T,r) = 1$ such that that the solution to equation (3.17) is given as

$$P(r,t;T) = e^{A(t;T)-B(t;T)\times r} \tag{4.44}$$

where

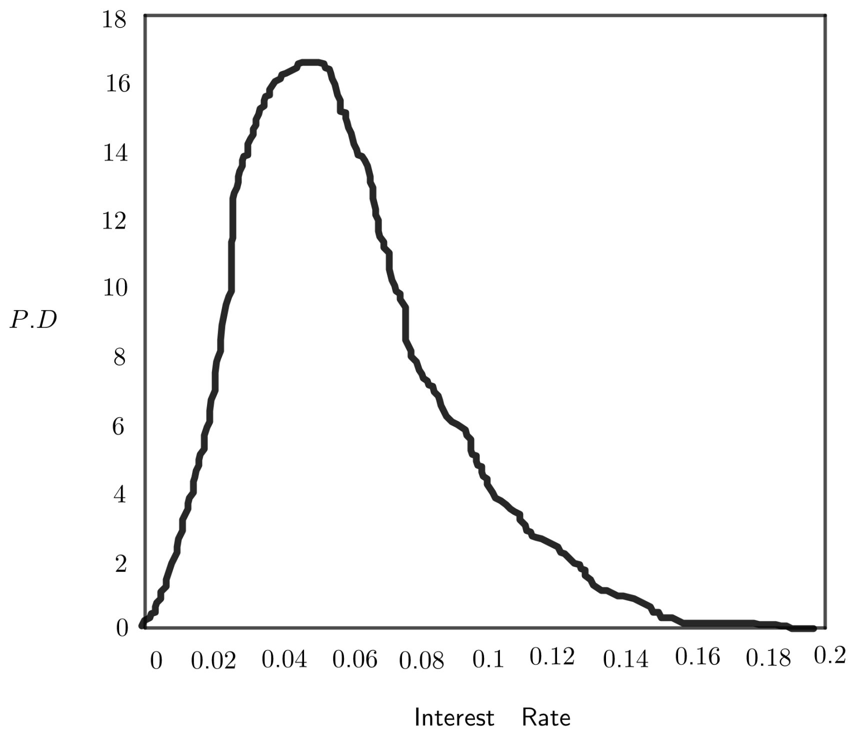
$$B(t;T) = \frac{2(e^{\nu_1(T-t)}-1)}{(\eta^*+\nu_1)(e^{\nu_1(T-t)}-1)+2\nu_1} \tag{4.45}$$

$$A(t;T) = \frac{2\xi^*\eta^*}{1}\log\left(\frac{2\nu_1 e^{(\nu_1+\eta^*)\frac{(T-t)}{2}}}{(\eta^*+\nu_1)(e^{\nu_1(T-t)}-1)+2\nu_1}\right) \tag{4.46}$$

and

$$\nu_1 = \sqrt{(\eta^*)^2+2} \tag{4.47}$$

*Fig*3.5 *Estimated Stationary Distribution for CIR Model*



where P.D stands for Probability distribution.

4.2.2 The Term Structure of Interest Rates Under the CIR Model

The types of yield curves that are possible under the CIR model are similar to that of the Vasicek model (see figure 4.1 above). Similar to Vasicek, the model suggests an ascending term structure of interest rates when the current short-term rate r_0 is low. On the contrary, if the current short-term rate r_0 is high, then the model suggests a declining term structure. As with the Vasicek model, the whole term structure is perfectly correlated with the short-term interest rate [7,17].

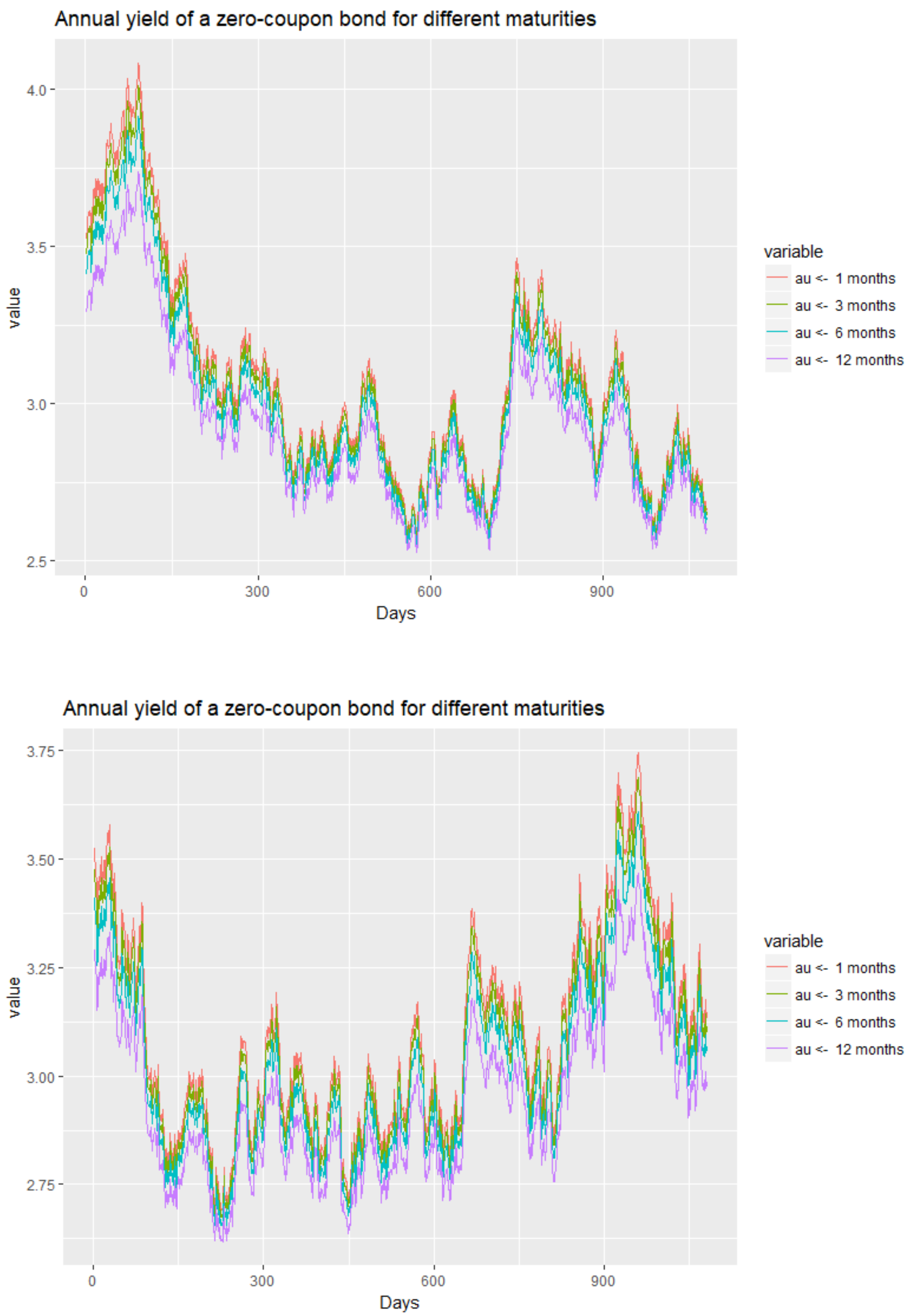

Annual yield of a zero-coupon bond for different maturities
value
4.0
3.5
3.0
2.5
0
300
600
900
Days
variable
au <- 1 months
au <- 3 months
au <- 6 months
au <- 12 months
Annual yield of a zero-coupon bond for different maturities
value
3.75
3.50
3.25
3.00
2.75
0
300
600
900
Days
variable
au <- 1 months
au <- 3 months
au <- 6 months
au <- 12 months

Figure 4.6:

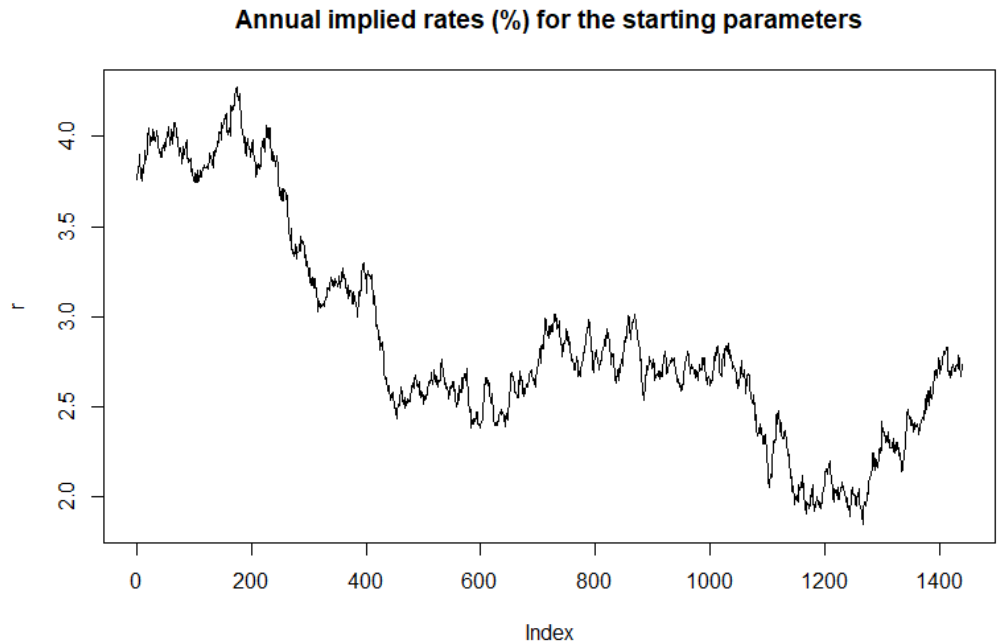


Figure 4.7:

The majority of models including the Vasicek model are partial equilibrium theories, since they take as input beliefs about future realizations of the short term interest rate and make assumptions about investors' preferences. The advantage of the equilibrium framework is that the term structure and its movements are determined endogenously by means of of the imposed equilibrium [3].

Chapter FIVE

Yield Curve Fitting and Arbitrage-Free Interest Rate Models

The previously discussed equilibrium models do not automatically fit today's term structure. Such a disadvantage could be remedied by altering parameters, such as taking parameters to be time dependent. By so doing, we are able to fit one or two bonds exactly. However, fitting the whole bonds which have different maturities leads to some errors.

To achieve this task of fitting to observed data, *No-Arbitrage Models* are designed to be exactly consistent with today's term structure of interest rates. In an equilibrium model, the term structure is an output, that is, by using such models, we estimate the value of the short rate for distinct time to maturity dates. However, the short rates for some time to maturity dates have been determined. Thus, there is a confliction between the model and the observed actual rate. Therefore, these kind of models leads to arbitrage. On the contrary, in no-arbitrage models, today's term structure is an input, that is, in the construction of these models, we take the observed actual rates to estimate the unobserved rates.

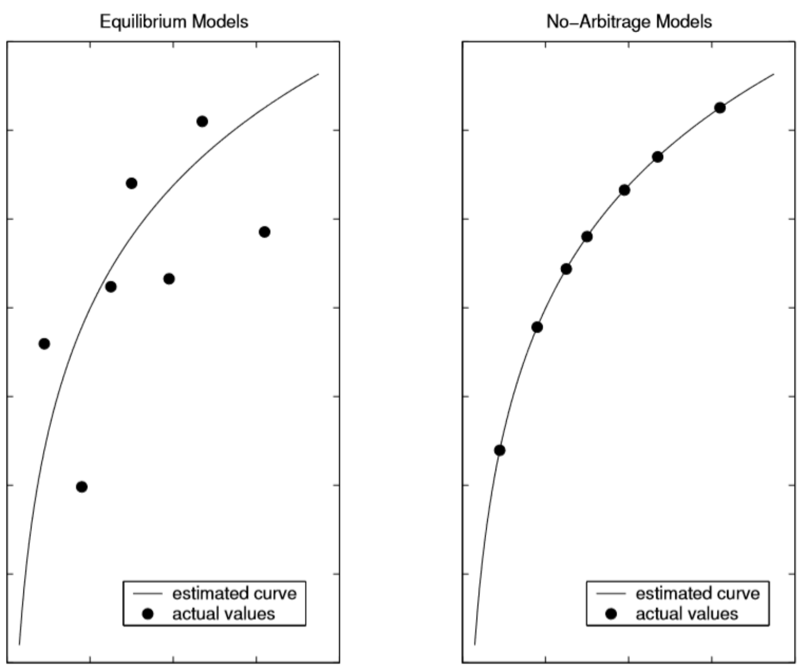


Figure 5.1: Estimation of the short rate with Equilibrium and the No-arbitrage Models

5.1 Ho-Lee Model

In 1986, Ho and Lee proposed the first no arbitrage model where the initial formulation is in the form of a binomial tree. The continuous time version of the model is given in the form

$$dr(t) = \xi(t)dt + \tau d\Phi(t) \tag{5.1}$$

where r is the short rate and τ is the constant instantaneous standard deviation of the short rate. The time dependent drift function $\xi(t)$ is chosen to ensure that the model fits the initial term structure. The main advantage of this approach is that it enables us to use the full information of the term structure to price contingent claims. An interesting fact of the model is that when it is used to price interest rate derivatives, the parameter that concerns the market price of risk proves to be irrelevant whereas the major drawback of the model is that it does not use the mean reversion property as in the case of the Vasicek model.

proof of the implicit solution

The Ho-Lee model is in the class of *affine term structure models* for which the value of the zero coupon bond price is given to be

$$P(t,T) = e^{a(t,T)+b(t,T)r_t} \tag{5.2}$$

We now determine $a(t,T)$ and $b(t,T)$ explicitly as follows.

For the market price of risk, we chose a constant q such that $q(r,T) = q$. With this, we have that $\xi(t) - q\tau = \xi(t)^*$ such that the fundamental pricing equation becomes

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial r}\xi(t)^* + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\tau^2 = rP. \tag{5.3}$$

To do this analytically, we look for solutions having the special form:

$$e^{a(t,T)+b(t,T)r_t} \tag{5.4}$$

To satisfy the boundary condition, we need to have

$$a(T,T) = b(T,T) = 0 \tag{5.5}$$

By substituting equation (5.4) in equation (5.3) we have

$$\left(\frac{\partial a}{\partial t} - r\frac{\partial b}{\partial t}\right)P - \xi(t)^* bP + \frac{\tau^2}{2}b^2 P = rP \tag{5.6}$$

or equivalently

$$\frac{\partial a}{\partial t} - \xi(t)^* b + \frac{\tau^2}{2}b^2 = r\left(1 + \frac{\partial b}{\partial t}\right) \tag{5.7}$$

We note that for fixed maturity T, a and b are functions of t only. Differentiating both sides of equations (5.7) with respect to r gives

$$0 = 1 + \frac{\partial b}{\partial t} \tag{5.8}$$

Substituting equation (5.8) in (5.7) we obtain

$$\frac{\partial a}{\partial t} - \xi(t)^* b + \frac{\tau^2}{2} b^2 = 0 \tag{5.9}$$

The general solution of (5.8) is computed as follows

$$1 + \frac{\partial b}{\partial t} = 0 \tag{5.10}$$

$$\Longrightarrow \int_t^T db(s,T) = -\int_t^T dt \tag{5.11}$$

$$\Longrightarrow b(T,T) - b(t,T) = -\int_t^T dt \tag{5.12}$$

$$\Longrightarrow b(t,T) = T - t \qquad \text{since} \quad b(T,T) = 0 \tag{5.13}$$

Using (5.13) and (5.9), $a(t,T)$ can be found by integration:

$$a(t,T) = \int_t^T \xi(s)^* b(s,T) ds - \frac{\tau^2}{2} \int_t^T b(s,t)^2 ds \tag{5.14}$$

$$\Longrightarrow a(t,T) = -\int_t^T \xi(s)^*(T-s) ds + \frac{\tau^2}{2} \int_t^T (T-t)^2 ds \tag{5.15}$$

$$\Longrightarrow a(t,T) = -\int_t^T \xi(s)^*(T-s) ds + \frac{\tau^2}{2} \Big[-\frac{1}{3}(T-s)^3 \Big]_t^T \tag{5.16}$$

$$\Longrightarrow a(t,T) = \int_t^T \xi(s)^*(s-T) ds + \frac{\tau^2}{2}\frac{1}{3}(T-t)^3 \tag{5.17}$$

5.2 The Hull-White Model

Up until 1990, the universally accepted equilibrium models (time-invariant models) such as the Vasicek model and the CIR model were probably considered relatively accurate around the time of their discovery. However, these time-invariant models are faced with the situation were the short term rate dynamics implies an internal term structure, which is not necessarily consistent with the observed term structure. That is, They do not provide a good fit to the initial term structure of interest rates, making them substandard for use by traders in pricing interest rates options [11, 12]. Practitioners are very unwilling to apply these models since they cannot be calibrated (i.e., forcing the model parameters to produce a curve as close as possible to the market curve) to actual yield curves. Moreover, these models cannot at the same time fit the initial term structure and a predefined future behavior for the short rate volatility.

For these reasons, Hull and White (1990) presented a class of models which enables both the initial fit of the term structure and a predefined future behavior for the short rate volatility under the assumption that the instantaneous short rate process evolves under the the risk neutral measure [11,19]. The most common description of the Hull and White (1993) is

$$dr(t) = (\xi(t) - \eta(t)r(t))dt + \tau(t)r^{\beta}(t)d\Phi(t) \tag{5.18}$$

with a defined risk premium

$$\lambda(r,t) = \lambda r^{\gamma} \tag{5.19}$$

and with λ, $\gamma \geq 0$. The functions $\xi(t), \eta(t)$ and $\tau(t)$ are deterministic functions of time that can be used to exactly calibrate the model to present market prices. However, this exact calibration renders the bond and bond option prices not analytically obtainable.

In the quest to exactly fit the initial term structure, it becomes very appealing to set $\eta(t)$ and $\tau(t)$ as time-varying parameters. A major consequence of such a decision is that the resulting volatility term structure will generally be non stationary and will progress in a quite uncertain (not predictable) way. Consequently, very fluctuating values from the parameters can often lead to a mis-specified or a mis-estimated model. Hull and White themselves stated the following *"It is always dangerous to use time-varying model parameters so that the initial volatility curve is fitted exactly. Using all the degrees of freedom in a model to fit the volatility exactly constitute an **over-parameterization** of the model. It is our option that there should be no more than one time varying parameter used in Markov models of the term structure evolution, and this should be used to fit the initial term structure".*

This explains why in practice, the model (5.18) is often applied with $\eta(t)$ and $\tau(t)$ constant and $\xi(t)$ as time-varying.

A major drawback of this model is the possibility of negative interest rates. However, since interest rates are becoming negative, this models could well be a point of departure of solving the problems those interest rates give rise to. Hull and White therefore present an extension of the Vasicel model[19].

5.2.1 The Hull-White Extended Vasicek Model

Hull and White (1990) suggested an extension of the Vasicek model (also known as $G++$ model) so that it can be consistent with both the present term structure of spot interest rates and the present term structure of interest rate volatility. According to the above Hull and White construction (equation (5.18)), the extended Vasicek model can be written as equation (4.1) with $\beta = 0$ as

$$dr(t) = \big(\xi(t) - \eta r(t)\big)dt + \tau_r d\Phi(t) \tag{5.20}$$

or alternatively as

$$dr(t) = \eta\bigg(\frac{\xi(t)}{\eta} - r(t)\bigg)dt + \tau_r d\Phi(t) \tag{5.21}$$

where η and τ are the positive constant mean reversion and volatility parameters respectively with $\Phi(t)$ a Wiener process. Again, $\frac{\xi(t)}{\eta}$ is the long term level to which the spot rate, $r(t)$, is moving, η is the rate at which the spot rate is pushing towards the long term level (i.e., an Ornstein-Urlenbeck process) [19]. The time dependent mean parameter $\xi(t)$ is chosen so that the model fits the initial term structure of interest rates presently observed in the market. The exact fitting of the term structure ensured by this model specifications makes it an *arbitrage-free model.* This model implies a normal distribution for the short interest rate process at each point (the reason for the possibility of negative interest rates). Furthermore, the model is analytical tractable and thus, the bonds can be priced easily.

In the sequel, we will try to determine the $\xi(t)$ that ensures an exact fitting of the model to the current observed term structure and subsequently solve the short rate dynamics of equation (5.20) explicitly. In that regards, let construct a money market economy by using the spot rate defined by equation (5.20) as

$$D(t) = e^{-\int_0^t r_u du} \tag{5.22}$$

$$dD(t) = -r_t D(t)dt = -r_t e^{-\int_0^t r_u du} dt \tag{5.23}$$

The formulation of the spot rate in equation (5.20) means that the Hull-White model belongs to the *affine class* of interest rate models so the prices of zero coupon bonds at time t for the time T maturity have the form as seen earlier (equation (5.2)) and at time t, the T-yield $y(t,T)$ is defined as

$$y(t,T) = -\frac{\ln P(t,T)}{T-t} \tag{5.24}$$

Our objective is to calibrate the time-varying mean drift $\xi(t)$ to the market today, that is, ensuring that the model prices of zero coupon bonds today, $P(0,T)$, fits (is equal) to the prices observed in the market.

5.2.2 Closed Form Solution for Prices of Zero Coupon Bonds

We want to explicitly find formulas for the functions $a(t,T)$ and $b(t,T)$ of equation (5.2) and hence the closed form solutions for zero coupon bonds in the Hull-White model [3,6].

For this purpose, we first proceed with the derivation of of the fundamental partial differential equation for zero coupon prices in the Hull-White model. To begin with, we find the dynamics of zero coupon bond prices by the use of Ito's lemma.

$$dP(t,T) = \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial r}dr(t) + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}(dr(t))^2 \tag{5.25}$$

By now incorporating equation (2.3) we have

$$\frac{dP(t,T)}{P(t,T)} = \left(\frac{\partial a(t,T)}{\partial t} + \frac{\partial b(t,T)}{\partial t}r(t)\right)dt + b(t,T)dr(t) + \frac{1}{2}b^2(t,T)(dr(t))^2$$

$$\frac{dP(t,T)}{P(t,T)} = \left(\frac{\partial a(t,T)}{\partial t} + \frac{\partial b(t,T)}{\partial t}r(t) + b(t,T)\xi(t) - \eta b(t)r_t + \frac{\tau^2 b^2}{2}(t,T)\right)dt + \tau b(t,T)d\Phi(t) \tag{5.26}$$

By applying the dynamics of the money market account given in equation (5.26) and the dynamics of the zero coupon bonds in equation (2.10), we find the dynamics of deflated zero coupon bonds prices as

$$dD(t)dP(t,T) = D(t)dP(t,T) + P(t,T)dD(t,T) + dD(t)dP(t,T) \tag{5.27}$$

where $dD(t)dP(t,T) = 0$.

Again, by incorporating the dynamics of (2.6) we get

$$\frac{dD(t)dP(t,T)}{D(t)P(t,T)} = \left(\frac{\partial a(t,T)}{\partial t} + \frac{\partial b(t,T)}{\partial t}r(t) + b(t,T)\xi(t) - \eta b(t)r_t + \frac{\tau^2 b^2}{2}(t,T) - r_t\right)dt + \tau b(t,T)d\Phi(t) \tag{5.28}$$

Since deflated prices are martingales under the equivalent martingale measure $\mathcal{Q}$ and according to the martingale presentation theorem, we have that the $dt-$term equal to zero and this holds for all t and r_t. Thus

$$\frac{\partial a(t,T)}{\partial t} + b(t,T)\xi(t) + \frac{\tau^2 b^2(t,T)}{2} = 0 \tag{5.29}$$

$$a(T,T) = 0 \tag{5.30}$$

$$\frac{\partial b(t,T)}{\partial t} - \eta b(t,T) - 1 = 0 \tag{5.31}$$

$$b(T,T) = 0 \tag{5.32}$$

solve equation 2.15

and thus the solution to the O.D.E is

$$b(t,T) = \frac{1}{\eta}\left(e^{-\eta(T-t)} - 1\right) \tag{5.33}$$

We can see that the derivative of $a(t,T)$ depends only on $b(t,T)$, so integrating $\frac{a(u,T)}{\partial u}$ on the closed interval $[t,T]$ give the solution to equation (2.13). So

$$\int_t^T \frac{a(u,T)}{\partial u} = a(T,T) - a(t,T) \tag{5.34}$$

and thus we have that

$$a(t,T) = \int_t^T b(u,T)\xi(u)du + \frac{1}{2}\int_t^T \tau^2 b^2(u,T)du \tag{5.35}$$

5.2.3 Calibration of the Current Yield Curve

To be able to match the model to current zero coupon prices, we need to chose $\xi(u)$ in equation (5.20) so that the initial yield curve is matched. For this purpose, we calibrate the model to the term structure of forward rates instead [6,20].

To begin with, let's recall that Forward rates are defined as

$$f^M(0,T) = -\frac{\partial \ln P(0,T)}{\partial T} = -\frac{\partial a(0,T)}{\partial T} - \frac{\partial b(0,T)}{\partial T}r_0 \tag{5.36}$$

From equation (5.33), we have that

$$\frac{\partial b(t,T)}{\partial T} = -e^{-\eta(T-t)} \tag{5.37}$$

Using Leibniz's rule for differentiating integral, we have from equations (2.14), (2.16) and (2.19) that

$$\frac{\partial a(0,T)}{\partial T} = b(T,T)\xi(T) + \int_0^T \frac{\partial b(u,T)}{\partial T}\xi(u)du - \frac{b^2(T,T)\tau^2}{2} + \tau^2 \int_0^T b(u,T)\frac{\partial b(u,T)}{\partial T}du \tag{5.38}$$

Inserting equations (2.17) and (2.21) in (2.22) yields

$$\frac{\partial a(0,T)}{\partial T} = -\int_0^T e^{-\eta(T-u)}\xi(u)du - \frac{\tau^2}{\eta}\int_0^T (e^{-\eta(T-u)} - 1)e^{-\eta(T-u)}du$$

$$\frac{\partial a(0,T)}{\partial T} = -\int_0^T e^{-\eta(T-u)}\xi(u)du - \frac{\tau^2}{\eta^2}\left[\frac{1}{2}(1 - e^{-2\eta T}) - (1 - e^{-\eta T})\right] \tag{5.39}$$

Combining everything together into equation (2.20), we get

$$f^M(0,T) = \int_0^T e^{-\eta(T-u)}\xi(u)du + \frac{\tau^2}{\eta^2}\left[\frac{1}{2}(1 - e^{-2\eta T}) - (1 - e^{-\eta T})\right] + e^{-\eta T} \tag{5.40}$$

Differentiating the above equation with respect to T, we have that

$$\frac{\partial f^M(0,T)}{\partial T} = \xi(T) - \eta\int_0^T e^{-\eta(T-u)}\xi(u)du + \frac{\tau^2}{\eta}\left(e^{-2\eta T} - e^{-\eta T}\right) - r_0\eta e^{-\eta T} \tag{5.41}$$

By using the first part of equation (2.24), equation (2.25) can be rewritten as

$$\frac{\partial f^M(0,T)}{\partial T} = \xi(T) - \eta f^M(0,T) + \frac{\tau^2}{\eta}\left[\frac{1}{2}(1 - e^{-2\eta T}) - (1 - e^{-\eta T})\right] + r_0\eta e^{-\eta T}$$
$$+ \frac{\tau^2}{\eta}\left(e^{-2\eta T} - e^{-\eta T}\right) - r_0\eta e^{-\eta T} \tag{5.42}$$

$$\frac{\partial f^M(0,T)}{\partial T} = \xi(T) - \eta f^M(0,T) - \frac{\tau^2}{2\eta}(1 - e^{-2\eta T}) \tag{5.43}$$

and thus

$$\xi(T) = \frac{\partial f^M(0,T)}{\partial T} + \eta f^M(0,T) + \frac{\tau^2}{2\eta}(1 - e^{-2\eta T}) \tag{5.44}$$

Now that we have gotten $\xi(T)$, we can now insert it in equation (2.19) to get the final expression of $a(t,T)$ That is

$$a(t,T) = \int_t^T b(u,T)\left[\frac{\partial f^M(0,u)}{\partial u} + \eta f^M(0,u) + \frac{\tau^2}{2\eta}(1 - e^{-2\eta u})\right]du + \frac{1}{2}\int_t^T \tau^2 b^2(u,T)du \tag{5.45}$$

We now compute the second integral in equation (2.19) as

$$\frac{\tau^2}{2}\int_t^T b(u,T)du = \frac{\tau^2}{2\eta^2}\int_t^T \left(e^{-\eta(T-u)} - 1\right)^2 du$$

$$\frac{\tau^2}{2}\int_t^T b(u,T)du = \frac{\tau^2}{2\eta^2}\left[\frac{1}{2\eta}(1 - e^{-2\eta(T-t)}) + (T-t) - \frac{2}{\eta}\left(1 - e^{-\eta(T-t)}\right)\right] \tag{5.46}$$

We next compute the first integral in equation (2.19) as

$$\int_t^T b(u,T)\xi(u)du = \frac{1}{\eta}\int_t^T (e^{-\eta(T-u)} - 1)\xi(u)du$$

$$\int_t^T b(u,T)\xi(u)du = \frac{1}{\eta}\int_t^T \xi(u)e^{-\eta(T-u)}du - \frac{1}{\eta}\int_t^T \xi du \tag{5.47}$$

Inserting equation (2.28) in the above equation

$$\int_t^T b(u,T)\xi(u)du = \frac{1}{\eta}\int_t^T e^{-\eta(T-u)}\left(\frac{\partial f^M(0,T)}{\partial T} + \eta f^M(0,T)\right)du$$
$$-\left(\frac{1}{\eta}\int_t^T \frac{\partial f^M(0,T)}{\partial T} - \eta f^M(0,T)\right)du + \frac{\tau^2}{2\eta}\int_t^T \big(e^{-\eta(T-u)} - 1\big)\big(1 - e^{-2\eta u}\big)du \quad (5.48)$$

where the last integral in the above expression of equation (2.32) is simplified as

$$\frac{\tau^2}{2\eta}\int_t^T \big(e^{-\eta(T-u)} - 1\big)\big(1 - e^{-2\eta u}\big)du = \frac{\tau^2}{2\eta^3}\left[1 - e^{-2\eta(T-t)} + \frac{e^{-2\eta T}}{2} - e^{-2\eta(T+t)} + \frac{e^{-2\eta T}}{2}\right]$$
$$+\frac{\tau^2(T-t)}{2\eta^2} \quad (5.49)$$

Thus

$$\int_t^T b(u,T)\xi(u)du = \frac{1}{\eta}\int_t^T e^{-\eta(T-u)}\frac{\partial f^M(0,T)}{\partial u}du + \int_t^T e^{-\eta(T-u)} f^M(0,u)$$
$$-\frac{1}{M}\Big[f^M(0,u)\Big]_t^T - \int_t^T f^M(0,u)du + \frac{\tau^2}{2\eta^3}\left[1 - e^{-2\eta(T-t)} + \frac{e^{-2\eta T}}{2} - e^{-2\eta(T+t)} + \frac{e^{-2\eta T}}{2}\right]$$
$$-\frac{\tau^2(T-t)}{2\eta^2} \quad (5.50)$$

Applying integration by part on the first term on the right hand side if equation (2.34)

$$\frac{1}{\eta}\int_t^T e^{-\eta(T-u)}\frac{\partial f^M(0,T)}{\partial u}du = \frac{1}{\eta}\Big[e^{-\eta(T-u)} f^M(0,T)\Big]_t^T - \int_t^T f^M(0,u)du \quad (5.51)$$

We thus have that

$$\int_t^T b(u,T)\xi(u)du = \frac{1}{\eta}\Big[e^{-\eta(T-u)} f^M(0,u)\Big]_t^T - \int_t^T e^{-\eta(T-u)} f^M(0,u)du + \int_t^T e^{-\eta(T-u)} f^M(0,u)du$$
$$-\frac{1}{\eta}\Big(f^M(0,T) - f^M(0,t)\Big) - \int_t^T f^M(0,u)du$$
$$+\frac{\tau^2}{2\eta^3}\left[1 - e^{-2\eta(T-t)} + \frac{e^{-2\eta T}}{2} - e^{-2\eta(T+t)} + \frac{e^{-2\eta T}}{2}\right] - \frac{\tau^2(T-t)}{2\eta^2} \quad (5.52)$$

Simplifying the above equation yields

$$\int_t^T b(u,T)\xi(u)du = -f^M(0,u)b(t,T) - \int_t^T f^M(0,u)du$$
$$+\frac{\tau^2}{2\eta^3}\left[1 - e^{-2\eta(T-t)} + \frac{e^{-2\eta T}}{2} - e^{-2\eta(T+t)} + \frac{e^{-2\eta T}}{2}\right] - \frac{\tau^2(T-t)}{2\eta^2} \quad (5.53)$$

Combining equations (2.33) and (2.37)

$$a(t,T) = \int_t^T b(u,T)\xi(u)du + \frac{1}{2}\int_t^T \tau^2 b^2(u,T)du$$
$$a(t,T) = -f^M(0,u)b(t,T) - \int_t^T f^M(0,u)du + \frac{\tau^2}{2\eta^3}\left[1 - e^{-2\eta(T-t)} + \frac{e^{-2\eta T}}{2} - e^{-2\eta(T+t)} + \frac{e^{-2\eta T}}{2}\right]$$
$$-\frac{\tau^2(T-t)}{2\eta^2} + \frac{\tau^2}{2\eta^2}\left[\frac{1}{2\eta}\big(1 - e^{-2(T-t)}\big) + (T-t) - \frac{2}{\eta}\big(1 - e^{-2(T-t)}\big)\right] \quad (5.54)$$

Again simplifying yields

$$a(t,T) = -f^M(0,u)b(t,T) - \int_t^T f^M(0,u)du + \frac{\tau^2 b^2(t,T)}{4\eta}\big(e^{-2\eta t} - 1\big) \quad (5.55)$$

and so we have that

$$\begin{aligned} P(0,T) &= e^{-\int_t^T f(0,u)}du \\ \implies \ln P(0,T) &= -\int_0^T f(0,u)du \qquad (5.56) \\ \implies \ln\left(\frac{P(0,T)}{P(0,t)}\right) &= -\int_t^T f(0,u)du \qquad (5.57) \end{aligned}$$

and hence

$$a(t,T) = -f^M(0,u)b(t,T) + \ln\left(\frac{P(0,T)}{P(0,t)}\right) + \frac{\tau^2 b^2(t,T)}{4\eta}\big(e^{-2\eta t} - 1\big) \qquad (5.58)$$

5.2.4 Closed Form Solution of the Hull-White Vasicek Model

To solve the S.D.E (2.3), we will employ the Ito's lemma on a defined function $g(r_t)$. Let's define this function as $g(r_t) = r_t e^{\eta t}$. Applying now Ito's lemma on $g(r_t)$

$$dg(r_t) = \xi(t)\eta e^{\eta t}dt + \tau e^{\eta t}d\Phi(t) \qquad (5.59)$$

Integrating both sides of equation (2.43) yields

$$\begin{aligned} \int_0^t dg(r_t) &= \eta\int_0^t \xi(s)e^{\eta s}ds + \tau\int_0^t e^{\eta s}d\Phi(s) \\ \implies g(r_t) - g(r_0) &= \eta\int_0^t \xi(s)e^{\eta s}ds + \tau\int_0^t e^{\eta s}d\Phi(s) \qquad (5.60) \end{aligned}$$

However, recall that $g(r_t) = r_t e^{\eta t} \implies r_t = g(r_t)e^{-\eta t}$ and thus $g(r_0) = r_0$. Hence, by multiplying both sides of equation (2.44) by $e^{-\eta t}$ and by performing the necessary re-arrangments, the solution of r_t is defined as

$$\begin{aligned} r_t e^{\eta s} - r_0 &= \eta\int_0^t \xi(s)e^{\eta s}ds + \eta\int_0^t e^{\eta s}d\Phi(s) \\ r_t &= r_0 e^{-\eta s} + \int_0^t \eta e^{-\eta t}e^{\eta s}\xi(s)ds + \tau\int_0^t e^{-\eta t}e^{\eta s}d\Phi(s) \\ &= r_0 e^{-\eta t} + \int_0^t e^{-\eta(t-s)}\xi(s)ds + \tau\int_0^t e^{-\eta(t-s)}d\Phi(s) \qquad (5.61) \end{aligned}$$

where equation (2.45) yields the short rate dynamics in level form as the solution to the S.D.E in (2.3). Now, since

$$E^{\mathcal{Q}}\big[\tau\int_0^t e^{-\eta(t-s)}d\Phi(s)|\mathcal{F}\big] = 0 \qquad (5.62)$$

then

$$E^{\mathcal{Q}}\big[r_t\big] = r_0 e^{-\eta t} + \int_0^t e^{-\eta(t-s)}\eta\xi(s)ds = r_0 e^{-\eta t} + \frac{\xi(s)}{\eta}e^{-\eta t}e^{-\eta s}\Big|_0^t = r_0 e^{-\eta t} + \frac{\xi(s)}{\eta}\Big[1 - e^{-\eta t}\Big] \qquad (5.63)$$

and

$$\lim_{t\to\infty} E^{\mathcal{Q}}\big[r_t\big] = \lim_{t\to\infty}\left[r_0 e^{-\eta t} + \frac{\xi(s)}{\eta}\left(1 - e^{-\eta t}\right)\right] = \frac{\xi(s)}{\eta} \qquad (5.64)$$

Also

$$Var_t^{\mathcal{Q}}\left[r_t\right] = \tau^2 \int_0^t e^{-2\eta(t-s)} ds = \frac{\tau^2}{2\eta} e^{-2\eta t} e^{-2\eta s}\Big|_0^t = \frac{\tau^2}{2\eta} e^{-2\eta t}\left[e^{2\eta t} - 1\right] = \frac{\tau^2}{2\eta}\left(1 - e^{-2\eta t}\right) \tag{5.65}$$

and

$$\lim_{t\to\infty} Var_t^{\mathcal{Q}}\left[r_t\right] = \lim_{t\to\infty}\left[\frac{\tau^2}{2\eta}\left(1 - e^{-2\eta t}\right)\right] = \frac{\tau^2}{2\eta} \tag{5.66}$$

Realize that equations (2.48) and (2.50) are also similar to those of the standard Vasicek model. Thus the distribution of $r(t)$ tends to $\mathcal{N}\left(\frac{\xi(s)}{\eta}, \frac{\tau^2}{2\eta}\right)$

5.2.5 Negative Rates

From the above analysis, we observe that the Hull-White model implies that the short rate has a normal distribution. So this short rate could technically take every value of $\mathbb{R}$, and a negative value [6,20]. We can compute the probability of this negative rate as

$$\begin{aligned} P(r(t) \leq 0) &= P\left(\sqrt{Var(r(t))}Z + E[r(t)] \leq 0\right) \\ &= P\left(Z \leq -\frac{E[r(t)]}{\sqrt{Var(r(t))}}\right) \end{aligned} \tag{5.67}$$

where Z is a random variable such that Z is standard normal between 0 and 1. Hence, we have

$$\begin{aligned} P(r(t) \leq 0) &= \phi\left(-\frac{E[r(t)]}{\sqrt{Var(r(t))}}\right) \\ &= \phi\left(-\frac{r_0 e^{-\eta t} + \int_0^t e^{-\eta(t-s)}\eta\xi(s)ds}{\sqrt{\tau^2\int_0^t e^{-2\eta(t-s)}ds}}\right) \\ &= \phi\left(-\frac{r_0 e^{-\eta t} + \frac{\xi(s)}{\eta}\left[1 - e^{-\eta t}\right]}{\sqrt{\frac{\tau^2}{2\eta}\left(1 - e^{-2\eta t}\right)}}\right) \end{aligned} \tag{5.68}$$

and in practical applications, this probability is often *"negligible"*. So when rates on the market are very low (approximately zero), the volatility tends to be also very low. This is equivalent (in Hull-White model) to to having a bigger mean reversion and a smaller $\xi(t)$.

Even though some times believed that having negative rates would be impossible in practice, it has however been recently seen that banks trading CHF (Swiss Franc) exchange a negative overnight rate.

5.2.6 The Hull-White Extended CIR Model

Hull and White also proposed an extension of the CIR model whose dynamics is given by

$$dr(t) = \big(\xi(t) - \eta(t)r(t)\big)dt + \tau(t)\sqrt{r(t)}d\Phi(t) \tag{5.69}$$

or alternatively as

$$dr(t) = \eta(t)\left(\frac{\xi(t)}{\eta(t)} - r(t)\right)dt + \tau(t)\sqrt{r(t)}d\Phi(t) \tag{5.70}$$

where $\xi(t)$, $\eta(t)$ and $\tau(t)$ are deterministic functions of time (i.e., non-random functions of time). Hull and White achieved this extension by inserting time varying parameter, ξ which is indeed the time varying mean [18]. Moreover, the process of matching the model and the market term structures of rates at the current time is achieved by solving a system with an infinite number of equations, one for each possible maturity. The diffusion term is usually taken to be constant in practice so as to ease calculations.

Furthermore, the analytic expression of $\xi(t)$ is not available and there is no guarantee that a numerical approximation of '$\xi(t)$' would keep the rate 'r' positive. Therefore, such an extension is not tractable analytically [22].

Never-the-less *Jamshidian* is 1995, proposed a simple version of equation (2.53) that turns out to be analytically tractable during which he assumed that for each t, the ratio $\frac{\xi(t)}{\tau^2(t)}$ equal a positive constant δ that must be greater than 0.5 to make sure that the origin is inaccessible [21].

Chapter SIX

Lognormal Models

Most of the models presented so far either have short rates or forward rates that are modeled as Gaussian processes. This reputation is due to the analytic tractability of Gaussian processes, but which also means that there is a possibility of a positive probability of negative rates leading to arbitrage opportunities in the presence of cash. This has led to a scramble of models capable of preventing negative rates. Authors therefore proposed models with lognormal rates in order to circumvent this problem [3,6,13]. Some of such models are presented below;

6.1 Black-Dermann-Troy (BDT) Model

In 1987, similar to the Ho-Lee model, BDT proposed a one factor binomial model whose continuous time equivalence is given to be

$$d\ln r = \big(\xi(t) - a\ln r\big)dt + \tau d\Phi \tag{6.1}$$

This model assumes a log-normal process for the short rate which makes it to avoid negative values [3]. However, in 1990, they extended the model to allow time dependent volatility:

$$d\ln r = \big(\xi(t) - a\ln r\big)dt + \tau(t)d\Phi \tag{6.2}$$

where $\xi(t)$ is chosen so that the model fits the term structure of short rates and $\tau(t)$ chosen to fit the term structure of short rate volatilities. Some of the reasons for which this model is so popular among practitioners are the following [6];

(1) It can be constructed to price exactly any set of discount bonds, as it uses the initially observed term structure to estimate the expected mean and standard deviations of future short rates,

(2) Swap rates which are a linear combination of discount bond, can be priced exactly for any volatility structure,

(3) Implied volatilities (caps or swaptions quotes) can be used directly to calibrate the model,

(4) The current market information can be represented by a simple recombining binomial tree with equally likely up and down moves, which ease computation and understanding.

Unfortunately, this model lacks analytical properties, and its implications and implicit assumptions are not known [6]. If the reversion rate and the volatility in the BDT model are decoupled, then we have

$$d\ln r = \big(\xi(t) - \eta(t)\ln r\big)dt + \tau(t)d\Phi \tag{6.3}$$

This new version is called the Black-Karasinski model.

6.2 Black Karasinski (BK) Model

In 1991, The Black and Karasinski proposed a binomial tree approach (discrete model) with time steps of varying lengths whose continuous time version is given in the sequel. BK model [16] is a one factor interest rate model that admits the specification of three time dependent factors and by so doing allowing the future short term interest rate volatilities to be specified independently of the initial volatility term structure. As in the BDT model, the short term interest rate is assumed to have a lognormal distribution at any time. The BK model is of the form

$$d\ln r(t) = \big(\xi(t) - \eta(t)\ln r(t)\big)dt + \tau(t)d\Phi(t) \tag{6.4}$$

with $d\Phi$ being the increment of a Wiener process, $\tau(t)$ being the volatility of the lognormal short rate process, $\eta(t)$ the mean reversion speed and $\xi(t)$ a constant function which allows the BK model to fit the actual yield curve.

The no-arbitrage arguments allow an equivalent formulation by mean of a parabolic PDE (which is backward in time) for the value V of a bond or a derivative security

$$\frac{\partial V}{\partial t} + \frac{1}{2}\tau^2 r^2(t)\frac{\partial^2 V}{\partial r^2(t)} + \left(\xi(t) + \frac{1}{2}\tau^2 - \eta(t)\ln r(t)\right)r(t)\frac{\partial V}{\partial r(t)} - r(t)V = 0 \tag{6.5}$$

Appropriate end conditions (payoff at maturity) and boundary conditions have to be formulated to make the pricing problem uniquely solvable. The equivalent formulation of the PDE in logarithm variable is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\tau^2 r^2(t)\frac{\partial^2 V}{\partial(\ln r(t))^2} + \left(\xi(t) - \eta(t)\ln r(t)\right)\frac{\partial V}{\partial(\ln r(t))} - e^{(\ln r(t))}V = 0 \tag{6.6}$$

One has to identify the BK parameters function $\xi(t)$, $\eta(t)$ and $\tau(t)$ form market prices of liquid instruments in order to be consistent with the market. Letting ξ to be the only time dependent function enables us to exactly fit the current term structure of interest rates and to keep the other two parameters at our disposal for calibration purposes.

For computational purposes, let $H(t) = lnr(t)$ and define a deterministic function as

$$K(t) = \int_0^t \eta(s)ds$$

such that

$$\frac{dK(t)}{dt} = \eta(t)$$

Now let $\Upsilon(t) = e^{K(t)}H(t)$.

$$\begin{aligned} d\Upsilon(t) &= e^{K(t)}\frac{dK(t)}{dt}H(t)dt + e^{K(t)}dH(t) \\ &= e^{K(t)}\eta(t)H(t)dt + e^{K(t)}\big[\big(\xi(t) - \eta(t)H(t)\big)dt + \tau(t)d\Phi\big] \\ &= e^{K(t)}\Big(\xi(t) + \tau(t)d\Phi(t)\Big) \end{aligned} \tag{6.7}$$

Integrating both sides of equation (4.7) we have

$$\begin{aligned} e^{K(t)}H(t) &= H(0) + \int_0^t e^{K(s)}\xi(s)ds + \int_0^t e^{K(s)}\tau(t)d\Phi(s) \\ H(t) &= e^{-K(t)}H(0) + \int_0^t e^{-(K(t)-K(s))}\xi(s)ds + \int_0^t e^{-(K(t)-K(s))}\tau(t)d\Phi(s) \end{aligned} \tag{6.8}$$

So for $u \leq t$,

$$H(t) = e^{-(K(t)-K(u))}H(u) + \int_u^t e^{-(K(t)-K(s))}\xi(s)ds + \int_u^t e^{-(K(t)-K(s))}\tau(t)d\Phi(s) \tag{6.9}$$

Replacing $\ln r(t)$ with $H(t)$ in equation (4.9)

$$\ln r(t) = e^{-(K(t)-K(u))}\ln r(u) + \int_u^t e^{-(K(t)-K(s))}\xi(s)ds + \int_u^t e^{-(K(t)-K(s))}\tau(t)d\Phi(s) \tag{6.10}$$

Preferably, we always want a process for the short term interest rate such that negative interest rates are avoided. The table below present some other one factor models.

Models	Equation for the short rate	
Merton	$dr = \xi dt + \tau dt\Phi$	
Dothan	$\tau r d\Phi$	
Geometrical Brownian motion	$\eta r dt + \tau r d\Phi$	
Brennan-Schwartz	$\big(\xi + \eta r\big)dt + \tau r d\Phi$	
Constant elasticity of variance	$\eta r dt + \tau r^{\gamma} d\Phi$	
Exponential Vasicek	$r(t) = e^{z(t)}, dz(t) = \eta\big(\xi - z(t)\big)dt + \tau d\Phi(t)$	

Table 1: One factor models.

Implicit within one factor models is the assumption that all information about future interest rates is contained in the current instantaneous short term interest rate and hence the price of all default free bonds may be represented as functions of this instantaneous rate and time only. Moreover, the instantaneous returns on bonds of all maturities are perfectly correlated. These features are inconsistent with reality and thus motivate the development of multi-factor models.

These single factor models have as main advantage that they are allowed to describe the short rate as a solution of a stochastic differential equation and through Markov theory use the associated partial differential equation. One major drawback is the fact that they cannot replicate adequately the market curve which depends on all the rate and not only on the short rate considered to build these models [3]. An alternative to short rate models is the Heath-Jarrow-Morton framework, capable of incorporating the entire dynamics of the yield curve.

Chapter SEVEN

Heath-Jarrow-Morton Framework

The vast majority of early interest rate models anterior to the Heath, Jarrow and Morton were some sought of extensions of the Vasicek model in which the spot rate was assumed to follow an Ornstein-Uhlenbeck process (mean reverting process) with constant volatility and constant mean reversion level. These models had as driving quantity the instantaneous spot rate of interest and were finite dimensional Markovian systems which allowed the standard no-arbitrage argument similar to that of the Black-Scholes and Merton. This then made possible the development of the pricing partial differential equation for the bond and bond option prices, and subsequently permitted the application of well established techniques from the theory of partial differential equations get analytical solutions, and numerical solutions in cases where analytical solutions were not possible [27].

These models were advantageous from the perspective that analytical solutions were usually obtainable but did not provide a good fit to the initial term structure of interest rates, so the calibration of model parameters to observed market data was not an easy task.

Contrary to these models was the framework introduced by Heath, Jarrow and Morton in 1992. This approach was a very general interest rate framework that took instantaneous forward rates as quantities driving the model, capable of including most, if not all, of the observable market features and had the useful feature that the model is automatically calibrated to the initial yield curve. This was an arbitrage free model for stochastic evolution of the entire term structure of interest rates, where the forward rates were determined by their instantaneous volatility structures. They choose the entire forward rate cure in their model which makes it a non-Markov process [28,29].

In its generality, the stochastic process for $f(t,T)$ in this model is assumed to be

$$f(t,T) = f(0,T) + \int_0^t \alpha_f(u,T)du + \sum_{i=1}^{n} \int_0^t \sigma_f^i(u,T)d\Phi_i(u) \tag{7.1}$$

where $f(t,T)$ is the forward rate at time t relevant to time $T > t$, $f(0,T)$ is the known initial forward rate curve, $\alpha_f(u,T)$ is the instantaneous forward rate's drift, $\sigma_f^i(u,T)$ are the volatilities of the forward rates, and $d\Phi_i$ is the i^{th} Wiener process (noise term). The forward rate process begins with initial value $f(0,T)$, and then progresses under a drift and several Wiener processes. There are n Wiener processes determining the stochastic fluctuation of the forward curve and they are all independent. $\alpha_f(T,T)$ and $\sigma_f^i(u,T)$ are adapted processes.

However, this interest rate dynamics does not necessarily provide a situation without arbitrage opportunities. To obtain an arbitrage free situation (that is, to obtain a unique equivalent martingale measure), Heath, Jarrow and Morton modeled the continuously compounded instantaneous forward rate $f(t,T)$, using the basic arbitrage relationship

$$f(t,T) = \frac{\partial \log p(t,T)}{\partial T} \tag{7.2}$$

from which

$$p(t,T) = e^{-\int_t^T f(t,s)ds} \tag{7.3}$$

They found out that for the dynamics to be arbitrage free, it is necessary that the drift $\alpha_f(u,T)$ should be associated to the volatility [29], so that the integrated dynamics of the instantaneous forward rate under the risk-neutral measure is

$$f(t,T) = f(0,T) + \int_0^t \sigma_f^i(u,T)\int_u^T \sigma_f^i(u,s)dsdu + \int_0^t \sigma_f^i(s,T)d\Phi_i(s) \tag{7.4}$$

At time $T = t$, the instantaneous short rate does not need to be modeled with a diffusion process but can be derived from the instantaneous forward rate as

$$r(t) = f(t,t) = f(0,T) + \int_0^t \sigma_f^i(u,t)\int_u^t \sigma_f^i(u,s)dsdu + \int_0^t \sigma_f^i(s,t)d\Phi_i(s) \tag{7.5}$$

However, this instantaneous short rate does not satisfy the Markov property but is is possible to spot particular volatilities that will render it a Markov process.
Carverhill, among others like Ritchken and Sankarasubramanian [30,31], discovered a special vector volatility of the type

$$\sigma_f^i(t,T) = R(t)S(T) \tag{7.6}$$

where R and S are strictly positive and deterministic functions of time. As a result, the short rate dynamics got the form

$$r(t) = f(t,t) = f(0,T) + S(t)\int_0^t R^2(u)\int_u^t S(s)dsdu + S(t)\int_0^t R(s)d\Phi_i(s) \tag{7.7}$$

7.1 Relation Between the HJM Framework with Short Rate Models.

Under the HJM framework, it is possible to get classes of interest rate models and every short rate model can be analogously stated in term of the forward rate theoretically[32]. Below are some short rate models under this framework.

7.1.1 Hull-White Extented Vasicek Method

Let's recall the Markovian dynamics of the short rate proposed by Carverhill (equation (5.7)). By considering only a single factor [32], it is possible to define deterministic and differential function V by

$$V(t) = f(0,t) + S(t)\int_0^t R^2(u)\int_u^t S(s)dsdu \tag{7.8}$$

with the aim to obtain

$$dr(t) = V^{'}(t) + S^{'}(t)\int_0^t R(u)d\Phi(s) + S(t)R(t)d\Phi(t)$$
$$\Longrightarrow\ dr(t) = \left[V^{'}(t) + S^{'}(t)\frac{r(t)-V(t)}{S(t)}\right]dt + S(t)R(t)d\Phi(t) \tag{7.9}$$

Now define

$$\eta(t) = -\frac{S^{'}(t)}{S(t)}, \qquad \xi(t) = V^{'}(t) - \frac{S^{'}(t)}{S(t)}V(t), \qquad \tau(t) = S(t)R(t) \tag{7.10}$$

With this specifications, the model then reduces to the Hull-White model

$$dr(t) = \big[\xi(t) - \eta(t)r(t)\big]dt + \tau(t)d\Phi(t) \tag{7.11}$$

where $\xi(t) = \eta(t)\theta(t)$. Moreover, by considering $\theta(t)$ as a constant coupled with trivial definition of the variables, we obtain the Vasicek model.

7.1.2 Hull-White Extented CIR Method

To get the dynamics of the short rate in the Hull-White extended CIR Under the HJM framework, it is important to define the volatility function as [32]

$$\sigma(s,t) = \nu(s)\sqrt{r(s)}e^{-\int_s^t k(v)dv} \tag{7.12}$$

such that

$$dr(t) = \left[\frac{\partial}{\partial t}f(0,t) + \eta(t)\big[f(0,t) - r(t)\big] + \mu(t)\right]dt + \nu(t)\sqrt{r(t)}d\Phi(t) \tag{7.13}$$

Finally, be defining

$$\xi(t) = \frac{\partial}{\partial t}f(0,t) + \eta(t)f(0,t) + \mu(t), \qquad \nu(t) = \tau(t), \qquad \xi(t) = \eta(t)\theta(t) \tag{7.14}$$

we obtain the Hull-White extended CIR method

$$dr(t) = \big(\xi(t) - \eta(t)r(t)\big)dt + \tau(t)\sqrt{r(t)}d\Phi(t) \tag{7.15}$$

Again, by defining some variables as constants instead of functions of time, we can obtain the CIR model.

7.1.3 Ho-Lee Model

In this case the volatility σ is considered to be constant such that the forward rate is given as

$$f(t,T) = f(0,T) + \int_0^t \sigma^2(T-u)du + \int_0^t \sigma d\Phi(s)$$
$$\Longrightarrow\ f(t,T) = f(0,T) + \sigma^2 t(T - \frac{t}{2}) + \sigma\Phi(t) \tag{7.16}$$

and in particular,

$$r(t) = f(t,t) = f(0,T) + \sigma^2 t\frac{t}{2} + \sigma\Phi(t) \tag{7.17}$$

Thus, by differentiating on both sides of the above equation with respect to t, the dynamics becomes

$$dr(t) = \left[\frac{\partial}{\partial t} f(0,t) + \sigma^2 t\right] dt + \sigma d\Phi(t) \tag{7.18}$$

Finally, by defining

$$\xi = \frac{\partial}{\partial t} f(0,t) + \sigma^2 t, \qquad \tau = \sigma \tag{7.19}$$

we obtain the Ho-Lee model

$$dr(t) = \xi(t)dt + \tau d\Phi(t) \tag{7.20}$$

7.1.4 How to Use HJM

In this section, we outline step-by-step how to use the HJM framework to determine prices of derivatives through numerical methods such as Monte carlo [32]

(i) Firstly, define the volatility structure $\sigma(t,T)$ by using one of the available model in the literature and then note the market instantaneous forward curve.

(ii) Secondly, simulate the growth of the entire forward rate curve in the risk neutral world until the desired date

$$df(t,T) = \sigma(t,T) \int_t^T \sigma(t,s)ds + \sigma(t,T)d\Phi(t)$$

(iii) Thirdly, Compute bond pricing for all dates using the formula

$$P(t,T) = e^{-\int_t^T f(t,s)ds} \tag{7.21}$$

(iv) Then, obtain the cash flows using forward rates.

(v) Next, consider the short rate in order to calculate the present value of the cash flows.

(vi) Finally, repeat step 2 in order to make enough realizations to have the discounted expected value with the desired precision.

The HJM framework, although broadly accepted as the most general and consistent framework under which interest rate derivatives can be studied, the major disadvantage of the HJM framework is that it produces models that are non-Markovian in general and thus the techniques from the theory of partial differential equations do not apply any more. The Monte Carlo simulation is can often be time consuming is the only method of solution for the general HJM model [27,28]. This difficulty, coupled to the lack of efficient numerical techniques under the general HJM framework favored the earlier models retain their popularity among practitioners. However, with the coming of of advance computer technology, the various forms of this framework are becoming popular once again among practitioners.

Bibliography

[1] Oldrich Vasicek. An equilibrium characterisation of the term structure. Journal of Financial Economics, 15, 1977. 45

[2] F. Jamshidian. An exact bond option pricing formula. Journal of Finance, 44:205209, 1989. 45, 48.

[3] Svoboda, Simona. 2004. Interest Rate Modelling. Palgrave Macmillan.

[4] Rogemar S. Mamon. Three Ways to Solve for the Bond Prices in the Vasicel Model. JOURNAL OF APPLIED MATHEMATICS AND DECISION SCIENCES, 8(1), 114 Copyright c 2004, Lawrence Erlbaum Associates, Inc.

[5] Anders Baek Holmgaard. Pricing of Contingent Interest Rate Claims, Foundations and Application of the Hull-White Extended Vasicek Term Structure model, Copenhagen Business School, 2013.

[6] R. Gibson, F. S. Lhabitant, Talay. Modeling the Term Structure of Interest Rates: A Review of the Literature, 2001.

[7] Pietro Veronesi. Fixed Income Securities: Valuation, Risk, and Risk Management. University of Chicago, Wiley, 2010.

[8] Simona Svoboda. An Investigation of Interest Rate Models and Their Calibration in the South African Market. February 2002.

[9] A Theory of the Term Structure of Interest Rates. Econometrica 53 (1985), 385-407

[10] Brigo, D. and F. Mercurio (2001): Interest Rate Models Theory and Practice, Springer, rst edition.

[11] J.C. Hull and A.D. White. A note on the models of hull and white for pricing options on the term structure. The Journal of Fixed Income, 5(2):97102, 1995.

[12] D. Filipovic. Term-Structure Models - A graduate Course, volume 1. Springer Finance, 2009.

[13] Jesper Lund. Lecture notes - review of continuous-time term-structure models - part i and associated slides. 1998.

[14] L.B.G. Andersen and V.V. Piterbarg. Interest Rate Modeling: Foundations and Vanilla Models, volume 1. Atlantic Financial Press, 2010.

[15] Andrew Lesniewski. Interest rate and fx models - the forward curve. Feb 17 - 2011.

[16] Black F. and Karasinski P., Bond option pricing when short rates are lognormal, Financial Analysts Journal, July-August (1991), 52-59

[17] Cox C.C., Ingersoll E.J. and Ross A.S., A theory of the term structure of interest rates, Econometrica, 53 (1985), 385-407.

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

[19] Hull J. and White A., One-factor interest rate models and the valuation of interest-rate derivative securities, Journal of Financial and Quantative Anaysis, 28 (1993), 235-254.

[20] Hull J. and White A., (1990) volume 3, number 4, p. 573-592. Pricing Interest Rate Derivative Securities, the Review of Financial Studies.

[21] Bjrk, T., (1998). Arbitrage Theory in Continuous Time, Oxford University Press New York.

[22] Pintoux, C., Privault, N., (2009). The Dothan pricing model revisited, Preprint.

[23] D. Brigo and F. Mercurio. Interest Rate Models, Theory and Practice. Springer Finance. Springer, second edition, 2006.

[24] Oran Cheyette. Interest Rate Models. Fixed Income Research, BARRA, Inc 2002.

[25] Paul Wilmott. Paul Wilmott on Quantitative Finance. John Wiley and Sons Ltd., 2000. 42

[26] Riccardo Rebonato. Term Structure Models: A Review. QUARC (QUAntitative Reseach Centre)- Royal Bank of Scotland Oxford University - OCIAM, 2003.

[27] Carl Chiarella and Oh Kang Kwon. A Class of Heath-Jarrow-Morton Term Structure Models with Stochastic Volatility. School of Finance and Economics, University of Technology Sydney, PO Box 123, Broadway NSW 2007.

[28] E. Valero, M. Torrealba, L. Lacasa and F. Fraysse. Fast Resolution of a Single Factor Heath-Jarrow-Morton Model with Stochastic Volatility. Dpto. de Matematica Aplicada, ETSI Aeronauticos, Universidad Politecnica de Madrid, Spain. Grupo BBVA, Mercados Globales y Distribucion, Desarrollo de Neuvos Productos, 2017.

[29] David Heath, Robert Jarrow and Andrew Morton. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuations. Econometrica, Vol.60, No.1 (January, 1992), 77-105.

[30] A. Carverhill, When is Spot Rate Markovian?, Mathematical Finance 4 (1994), no. 4, 305312.

[31] P. Ritchken and L. Sankarasubramanian, Volatility Structures of Forward Rates and the Dynamics of the Term Structure, Mathematical Finance 5 (1995), no. 1, 5572.

[32] Luca Beltrami. A Calibration Method For the HJM Model Based on the Levenberg-Marquardt Optimaization Algorithm. Politecnico de Milano, 2009-2010.

[33] Bank for International Settlements, Outstanding World Bond Market Debt , 2009.

[34] L. C. G. Rogers. Which model for term structure of interest rates should one use? Mathematical Finance, IMA Volume 65 (1995) 93-116.

[35] Frank J. Fabozzi. The Hand Book of Fixed Income Securities. Eight edition, McGraw-Hill, 1983.

[36] Leslie Jane Federer Vaaler and James W. Daniel. Mathematical Interest Theory. Second edition, MAA, 2009.

[37] Robert L. McDonald. Derivative Markets, Third edition, PEARSON, 2014.