palgrave macmillan

Interest Rate Modelling

Simona Svoboda

Interest rate modelling

Simona Svoboda

palgrave
macmillan

First published 2004 by
PALGRAVE MACMILLAN
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N. Y. 10010
Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN 978-1-349-51732-9 ISBN 978-1-4039-4602-7 (eBook)
DOI 10.1057/9781403946027

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources.

A catalogue record for this book is available from the British Library.

A catalogue record for this book is available from the Library of Congress

10 9 8 7 6 5 4 3 2 1
13 12 11 10 09 08 07 06 05 04

Contents

Introduction

Growth in the derivatives markets has brought with it an ever-increasing volume and range of interest rate dependent derivative products. To allow profitable, efficient trading in these products, accurate and mathematically sound valuation techniques are required to make pricing, hedging and risk management of the resulting positions possible.

The value of vanilla European contingent claims such as caps, floors and swaptions depends only on the level of the yield curve. These types of instruments are priced correctly using the simple model developed by Black [**5**]. This model makes several simplifying assumptions which allow closed-form valuation formulae to be derived. This class of vanilla contingent claims has become known as 'first-generation' products.

These instruments expose investors to the level of the underlying yield curve at one point in time. They reflect the investors' view of the future changes in the level of the yield curve, not their view of changes in the slope of the curve. 'Second-' and 'third-generation' derivatives, such as path-dependent and barrier options, provide exposure to the relative levels and correlated movements of various portions of the yield curve. Rather than hedging these exotic options with the basic underlying instrument, i.e. the bond, the 'first generation' instruments are used. Therefore, the Black model prices of these 'first generation' instruments are taken as given. This does not necessarily imply a belief in the intrinsic correctness of the Black model. Distributional assumptions which are not included in the Black model, such as mean reversion and skewness, are incorporated by adjusting the implied volatility input.

The more sophisticated models developed allow the pricing of instruments dependent on the changing level and slope of the yield curve. A crucial factor is that these models must price the exotic derivatives in a manner that is consistent with the pricing of vanilla instruments. When assessing the correctness of any more sophisticated model, its ability to reproduce the Black prices of vanilla instruments is vital. It is not a model's *a priori* assumptions, but rather the correctness of its hedging performance that plays a pivotal role in its market acceptance.

The calibration of the model is an integral part of its specification, so the usefulness of a model cannot be assessed without considering the reliability and robustness of parameter estimation.

General framework

The pricing of interest rate contingent claims has two parts. Firstly, a finite number of pertinent economic fundamentals are used to price all default-free zero coupon (discount) bonds of varying maturities. This gives rise to an interest rate term structure, which attempts to explain the relative pricing of zero coupon bonds of various maturities. Secondly, taking these zero coupon bond prices as given, all interest rate derivatives may be priced.

As with asset prices, the movement of interest rates is assumed to be determined by a finite number of random shocks, which feed into the model through stochastic processes. Assuming continuous time and hence also continuous interest rates, these sources of randomness are modelled by Brownian motions (Wiener processes).

When modelling interest rates we do not have a finite set of assets, but rather a one-parameter family of assets: the discount bonds, with the maturity date as the parameter. The risk-free rate of interest (short-term interest rate) is not specified exogenously (as in stock price models), but is the rate of return on a discount bond with instantaneous maturity. Also, unlike in asset pricing theory, the fundamental assets – the discount bonds, may themselves be viewed as derivatives. Hence the modelling of the interest rate term structure may be viewed as tantamount to interest rate derivative pricing.

The theory of interest rate dynamics relies on a degree of abstraction in that the fundamental assets (the discount bonds) are assumed to be perfect assets, that is default-free and available in a continuum of maturities.

Approaches to term structure modelling

The modelling of the term structure of interest rates in continuous time lends itself to various approaches. The most widely used approach has been to assume the short-term interest rate follows a diffusion process[1]. Bond prices are then determined as solutions to a partial differential equation which places restrictions on the relationship of risk premia of bonds of varying maturities. Unfortunately it is particularly difficult and cumbersome to fit the observed term structure of interest rates and volatilities within this simple diffusion model.

One factor models. One of the first models to make a significant impact on interest rate modelling was by Vasicek [**50**]. Although his paper is entitled *"An Equilibrium Characterisation of the Term Structure"* he does not make any assumptions about equilibrium within an underlying economy, nor does he make use of an equilibrium argument in the derivation. Instead, his derivation

[1]A diffusion process is a Markovian process for which all realisations or sample functions $\{X_t, t \in [0,\infty)\}$ are continuous. A Markovian process has the characteristic that given the value of X_t, the values of $X_s, s > t$ do not depend on the values of $X_u, u < t$. Brownian motion is a diffusion process. [**34**]

relies on an arbitrage argument, much like the one used by Black and Scholes in the derivation of their option pricing model. Vasicek makes assumptions about the stochastic evolution of interest rates by exogenously specifying the process describing the short-term interest rate.

A later approach utilised by Cox, Ingersoll and Ross (CIR) [**18**] begins with a rigorous specification of an equilibrium economy which becomes the foundation for the model specifications. Assumptions are made about the stochastic evolution of exogenous state variables and about investor preferences. The form of the short-term interest rate process and hence the prices of contingent claims are endogenously derived from within the equilibrium economy. The CIR model is a complete equilibrium model, since bond prices are derived from exogenous specifications of the economy, that is: production opportunities, investors' tastes and beliefs about future states of the world.

Most models, including the Vasicek model, are partial equilibrium theories, since they take as input beliefs about future realisations of the short-term interest rate (depicted within the functional form of the short-term interest rate process) and make assumptions about investors' preferences (specified by the market prices of risk). The resulting discount bond yields are based on these assumptions.

The equilibrium approach has the advantage that the term structure, its dynamics and the functional form of the market prices of risk are endogenously determined by means of the imposed equilibrium. CIR [**18**] criticise the partial equilibrium approach, since it applies an arbitrage argument to exogenously specified interest rate dynamics and allows an arbitrary choice of the form of the market prices of risk which may lead to internal inconsistencies.

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

Multi-factor models. Brennan and Schwartz [**10**] propose an interest rate model based on the assumption that the whole term structure can be expressed as a function of the yields of the longest and shortest maturity default-free bonds. Longstaff and Schwartz [**38**] develop a two factor model of the term structure based on the framework of Cox, Ingersoll and Ross [**18**]. The two factors are the short-term interest rate and the instantaneous variance of changes in this short-term interest rate (volatility of the short-term interest rate). Therefore the prices of contingent claims reflect the current levels of the interest rate and its volatility. Langetieg [**36**] develops a general framework where the short-term interest rate is expressed as the sum of a number of underlying stochastic factors. The model is essentially an extension of Vasicek's approach where the evolution of the short-term interest rate is subject to multiple sources of uncertainty.

The use of two or more factors improves the explanatory power of the models, but increases the degree of numerical complexity. Identifying additional factors is quite difficult and cumbersome numerical procedures need to be used. Although not always explicitly stated, the multi-factor models rely on the assumption of market completeness. This means there must be at least as many tradable assets as there are sources of uncertainty. If this is not the case, the market is incomplete and there are stochastic fluctuations in the Brownian motions which are not picked up as price changes in some asset, and hence cannot be hedged.

An attempt at preference-free pricing. In the above models the underlying stochastic state variables are interest rates or other non-tradable securities. This means that the resulting valuation formulae for contingent claims depend on investor preferences, and empirical approximations must be used to estimate the investor-specific variables.

In an attempt to avoid this problem, Ball and Torous [**4**] propose a model where the underlying state variable is the bond directly. The price of a risk-free zero coupon bond is assumed to follow a Brownian bridge process. The specification of this process ensures that the price of the bond converges to its face value at maturity. Also, since this underlying state variable is a tradable security, a preference-free closed-form valuation formula for European options may be derived. However, this model has shortcomings which make it unsuitable.

Fitting the initial term structure. The above models attempt to model interest rates so as to produce a realistic future yield curve. No explicit attempts are made to match the current observed term structure. CIR [**18**] mention a possible extension to their model that allows a time-dependent drift term and point out that information contained within the initial term structure could be used to determine the drift without placing any restrictions on its functional form. This point was taken up much later by Hull and White (HW) [**29**] in their extension of the Vasicek and CIR models. Hull and White propose an extension to these models allowing time-dependent drift and volatility parameters. The extended Vasicek model allows analytical solutions for bonds and bond options. Model parameters, including those involving the market price of risk, are determined in terms of the initial term structure. This approach allows an exact fit to the initial term structure of interest rates and possibly also interest rate volatilities.

Black, Derman and Toy (BDT) [**6**] developed a one-factor discrete time model of the term structure. A binary tree of one period interest rates is constructed in such a way that the rate and transition probabilities at each node match an initial observed term structure of interest rates and volatilities. Here the one-period rate is the analogue of the short-term interest rate within a continuous time setting. This model is in fact a time-discretisation of a diffusion model where the short-term interest rate is lognormal.

Modelling the forward rate. Ho and Lee [**27**] introduced a new approach to term structure modelling. Instead of modelling the short-term interest rate, they developed a discrete time model of the evolution of the whole yield curve. The short-term interest rate is a single point on the yield curve, which, in one-factor models, is assumed to be the only factor determining the entire yield curve. The Ho and Lee model admits an arbitrary specification of the initial yield curve, so it may be calibrated to the observed initial yield curve.

Heath, Jarrow and Morton (HJM) [**25**] developed a general framework of interest rate dynamics allowing an arbitrary specification of the initial term structure. They approached the problem by exogenously specifying the dynamics of instantaneous, continuously compounded forward rates. Rather than a traditional no arbitrage argument (as used, for example, in the derivation of the Vasicek model) they use a change of probability measure technique initially formulated by Harrison and Kreps [**23**] and Harrison and Pliska [**24**]. This involves transforming to the risk-neutral measure under which all asset prices have the same drift, that is the risk-free rate of interest. Within a complete market, without profitable arbitrage opportunities, this risk-neutral measure is unique and allows all discounted[2] asset prices to be martingales. Default-free zero coupon bonds and derivative securities may now be valued by taking the expectation, under the risk-neutral measure, of the discounted terminal (maturity) value.

The Ho and Lee model is a special case of the HJM framework and may be viewed as its discrete time predecessor.

Outline of the book

Part examines the various models mentioned above. Each chapter begins with the assumptions underlying each model and examines its derivation and, where analytical solutions exist, the derivation of the pricing formulae for contingent claims. Comparisons are drawn between the various models with the aim of explaining the significance of the various approaches. Part 12.8 describes the calibration procedure for the HW-extended Vasicek, BDT and HJM models. These models represent distinct approaches to term structure modelling and their calibration methodologies are representative of the general class of models to which each belongs.

[2]Here, 'discounted' implies that the asset prices are expressed as a ratio of the money market account.

Part I: Interest Rate Models

CHAPTER 1

The Vasicek Model

The initial formulation of Vasicek's model is very general, with the short-term interest rate being described by a diffusion process. An arbitrage argument, similar to that used to derive the Black–Scholes option pricing formula [**8**], is applied within this broad framework to determine the partial differential equation satisfied by any contingent claim. A stochastic representation of the bond price results from the solution to this equation. Vasicek then allows more restrictive assumptions to formulate the specific model with which his name is associated.

The consistency of the model specifications with an underlying economic equilibrium is not proved. Rather, it is implicitly assumed. The special case of the general model formulation, which Vasicek uses for illustrative purposes, was suggested by Merton [**40**] in a study of price dynamics in a continuous time, equilibrium economy. Equilibrium conditions imply that interest rates are such that the demand and supply of capital are equally matched.

1.1. Preliminaries

First, define the following variables:

$P(t,T)$	–	time t price of a discount bond maturing at time T, $t \leq T$, with $P(T,T) = 1$.
$R(t,\tau)$	–	time t rate of interest applicable for period τ. In terms of the return on a discount bond, this rate is defined as the internal rate of return, at time t, on a bond with maturity date $T = t + \tau$.
$r(t)$	–	instantaneous rate of interest (short rate) at time t.
$f(t,T)$	–	instantaneous forward interest rate i.e. time t assessment of the instantaneous rate of interest applicable at time T.

The following relationships apply:

$$
\begin{aligned}
R(t,\tau) &= -\frac{1}{\tau}\log P(t,t+\tau), \quad \tau > 0 \\
R(t,\tau) &= \frac{1}{\tau}\int_t^{t+\tau} f(t,\tau)d\tau
\end{aligned}
\tag{1.1}
$$

or explicitly for the forward rate:

$$
f(t,T) = \frac{\partial}{\partial T}\left[(T-t)R(t,T-t)\right]
\tag{1.2}
$$

The short rate is defined as the instantaneous rate of interest at time t:

$$
\lim_{\tau\to 0} R(t,\tau) = R(t,0) = r(t)
\tag{1.3}
$$

Vasicek makes the following three assumptions:

Assumption 1. The current short interest rate is known with certainty. However, subsequent values of the short rate are not known. The assumption is made that $r(t)$ follows a stochastic process. Also assume that $r(t)$:

- is a continuous function of time,
- follows a Markovian process. That is, given its current value, future developments of the short rate are independent of past movements.

This implies that the short rate process is fully characterised by a single state variable, i.e. its current value, and the probability distribution of $r(t^*), t^* \geq t$ is fully determined by $r(t)$. A continuous Markovian process is called a diffusion process, which is described by the stochastic differential equation:

$$
dr = v(r,t)dt + s(r,t)dz
\tag{1.4}
$$

where $v(r,t)$ is the instantaneous drift and $s^2(r,t)$ the instantaneous variance of $r(t)$. $z(t)$ is a Wiener process under a given measure Q.

Assumption 2. The time t price of a discount bond with maturity T, $P(t,T)$, is fully determined by the time t assessment of $\{r(t^*), t \leq t^* \leq T\}$, the segment of the short rate over the remaining term of the bond. Moreover, the development of the short rate over $[t,T]$ is fully determined by its current value $r(t)$, so the bond price may be written as a function of the current short rate:

$$
P(t,T) = P(r(t),t,T)
\tag{1.5}
$$

Hence the entire term structure is determined by the short rate.

Assumption 3. The market is assumed to be efficient. This implies:

- there are no transaction costs;
- information is simultaneously distributed to all investors;
- investors are rational with homogeneous expectations;
- profitable, riskless arbitrage is not possible.

1.2. Term structure equation

Equation (1.5) implies that the bond price $P = P(r(t), t, T)$ is a function of the short rate. Applying Ito's Lemma to the bond price and using (1.4) we derive the stochastic differential equation for the bond price:

$$dP = \left(v(r,t)\,\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} + \frac{1}{2}s^2(r,t)\frac{\partial^2 P}{\partial r^2} \right) dt + s(r,t)\frac{\partial P}{\partial r} dz \tag{1.6}$$

Set

$$\mu(t,T) = \frac{1}{P}\left(v(r,t)\,\frac{\partial}{\partial r} + \frac{\partial}{\partial t} + \frac{1}{2}s^2(r,t)\,\frac{\partial^2}{\partial r^2} \right) P \tag{1.7}$$

$$\sigma(t,T) = -\frac{1}{P}\,s(r,t)\,\frac{\partial}{\partial r}P \tag{1.8}$$

and hence

$$dP = \mu(t,T)P\,dt - \sigma(t,T)P\,dz \tag{1.9}$$

where $\mu(t,T)$ and $\sigma^2(t,T)$ are the time t mean and variance of the instantaneous rate of return of a T-maturity zero coupon bond.

Since a single state variable is used to determine all bond prices, the instantaneous returns on bonds of varying maturities are perfectly correlated. Hence a portfolio of positions in two bonds with different maturity dates can be made instantaneously risk-free. This means that the instantaneous return on the portfolio will be the risk-free rate of interest. Consider a time t portfolio of a short position in V_1 bonds with maturity T_1 and a long position in V_2 bonds with maturity T_2. The change over time in the value of the portfolio, $V = V_2 - V_1$, is obtained from (1.9):

$$dV = (V_2\mu(t,T_2) - V_1\mu(t,T_1))\,dt - (V_2\sigma(t,T_2) - V_1\sigma(t,T_1))\,dz \tag{1.10}$$

Choosing V_1 and V_2 such that the coefficient of the Wiener coefficient in (1.10) reduces to zero will result in a portfolio with a strictly deterministic instantaneous return. Hence we require:

$$\begin{aligned} V_2\sigma(t,T_2) - V_1\sigma(t,T_1) &= 0 \\ (V + V_1)\sigma(t,T_2) - V_1\sigma(t,T_1) &= 0 \\ V_1\sigma(t,T_1) - V_1\sigma(t,T_2) &= V\sigma(t,T_2) \end{aligned}$$

$$\Rightarrow V_1 = \frac{V\sigma(t,T_2)}{\sigma(t,T_1) - \sigma(t,T_2)}$$

Similarly:

$$V_2 = \frac{V\sigma(t,T_1)}{\sigma(t,T_1) - \sigma(t,T_2)}$$

and hence (1.10) becomes:

$$dV = V\,\frac{\mu(t,T_2)\sigma(t,T_1) - \mu(t,T_1)\sigma(t,T_2)}{\sigma(t,T_1) - \sigma(t,T_2)}\,dt$$

Invoking assumption 3, that no riskless arbitrage is possible, the instantaneous return on the portfolio must be the risk-free rate, $r(t)$. That is:

$$\frac{\mu(t,T_2)\sigma(t,T_1)-\mu(t,T_1)\sigma(t,T_2)}{\sigma(t,T_1)-\sigma(t,T_2)}=r(t)$$

Rearranging the terms in the equation, we have:

$$\frac{\mu(t,T_1)-r(t)}{\sigma(t,T_1)}=\frac{\mu(t,T_2)-r(t)}{\sigma(t,T_2)} \tag{1.11}$$

and since this equality is independent of the bond maturity dates, T_1 and T_2, we can define:

$$q(r,t)=\frac{\mu(t,T)-r(t)}{\sigma(t,T)} \quad t\leq T \tag{1.12}$$

where $q(r,t)$ is independent of T. $q(r,t)$ measures the increase in expected instantaneous return on a bond, for a unit increase in risk, and is referred to as the market price of risk. Substituting the formulae for μ and σ from (1.7) and (1.8), we derive a partial differential equation for the bond price:

$$\frac{\partial P}{\partial t}+(v(r,t)+s(r,t)q(r,t))\frac{\partial P}{\partial r}+\tfrac{1}{2}s(r,t)^2\frac{\partial^2 P}{\partial r^2}-r(t)P=0, \quad t\leq T \tag{1.13}$$

This equation, referred to as the term structure equation, is a general zero coupon bond pricing equation in a market characterised by assumptions 1, 2 and 3. To solve (1.13), we need to specify the parameters of the short-term interest rate process defined by (1.4), the market price of risk $q(r,t)$, and apply the boundary condition:

$$P(r,T,T)=1 \tag{1.14}$$

Using equation (1.1) we can evaluate the entire term structure, $R(t,\tau)$.

1.3. Risk-neutral valuation

The mean rate of return on a bond can be written as a function of its variance, the risk-free interest rate and market price of risk. From (1.12) we have:

$$\mu(t,T)=r(t)+q(r,t)\,\sigma(t,T) \qquad \forall t\leq T$$

Hence the bond price dynamics (1.9), may be written in terms of the market price of risk as[1]:

$$\begin{aligned} dP &= (r+q\sigma)P\,dt-\sigma P dz \\ &= rP\,dt-\sigma P(-q\,dt+dz) \end{aligned}$$

Let $d\tilde{z}=-q\,dt+dz$ and the above equation becomes:

$$dP=rP\,dt-\sigma Pd\tilde{z} \tag{1.15}$$

[1]To lighten the notation, the functional dependence of r, q, v, s and σ on r, t and T is suppressed.

where $\tilde{z}$ is the Wiener process in the risk-neutral world being governed by the probability measure. Since $dz = q\,dt + d\tilde{z}$, the equation describing the dynamics of the short-term rate, (1.4) may be written in terms of $d\tilde{z}$ as:

$$\begin{aligned} dr &= vdt + s\,dz \\ &= vdt + s(q\,dt + d\tilde{z}) \\ &= (v + sq)\,dt + s\,d\tilde{z} \end{aligned} \tag{1.16}$$

Using the dynamics of the bond price in (1.15) and the PDE of the bond price (1.13) with boundary condition $P(T,T) = 1$, the Feynman–Kac theorem[2] may be applied to yield the valuation:

$$P(t,T) = \widetilde{\mathbb{E}}_t \left[\exp\Big(-\int_t^T r(u)\,du \Big) P(T,T) \right] \tag{1.17}$$

Here we take the expectation with respect to $\widetilde{\mathbb{E}}$ which corresponds to the risk-neutral world. $\widetilde{\mathbb{E}}$ corresponds to the equivalent probability measure $\widetilde{Q}$ which utilises risk-neutral probabilities. (As opposed to the utility dependent probability measure, Q, which represents investor specific probabilities.) By introducing the market price of risk q we are able to transform the probability measure from a utility-dependent to a risk-neutral one. The Girsanov Theorem[3] defines this transformation. First consider the Wiener process:

$$\tilde{z} = -\int_0^t q\,du + z$$

Let $t \le t^* \le T$ and define

$$Z(t^*) = \exp\Big(\int_t^{t^*} q\ dz - \frac{1}{2} \int_t^{t^*} q^2 du \Big)$$

as the Radon–Nikodym derivative used to define the new probability measure, that is:

$$\frac{d\widetilde{Q}}{dQ} = Z(T)$$

[2]The discounted Feynman–Kac theorem is applicable in this case. This theorem defines the relationship between a stochastic differential equation (SDE) and the corresponding partial differential equation (PDE). Considering the SDE:

$$dX(s) = \alpha(X(s))ds + \gamma(X(s))dz(s)$$

Let $0 \le t \le T$ where $T > 0$ is fixed, and let $h(y)$ be some function. Define:

$$\nu(t,x) = \mathbb{E}\left[e^{-r(T-t)} h(X(T)) \Big| X(t) = x \right]$$

Then the corresponding PDE is:

$$\nu_t(t,x) + \alpha(x)\nu_x(t,x) + \tfrac{1}{2}\gamma^2(x)\nu_{xx}(t,x) = r\nu(t,x)$$

See [**45**] for more details.

[3]For more details about the application of Girsanov's Theorem and the Radon–Nikodym derivative in the change of measure see [**45**] and [**41**].

Also, expectation with respect to $\widetilde{\mathbb{E}}$ is calculated as:

$$\widetilde{\mathbb{E}}\,[Y] = \mathbb{E}\,[Z.Y]$$

for any random variable Y. Hence the expected bond price (1.17) may be expressed in terms of the utility-dependent measure as:

$$P(t,T) = \mathbb{E}_t\left[\exp\Big(-\int_t^T r(u)\,du - \frac{1}{2}\int_t^T q^2\,du + \int_t^T q\,dz\Big)P(T,T)\right] \tag{1.18}$$

1.4. Note on empirical estimation of the market risk premium

Empirical testing and application of this model requires the specification of three parameters: the instantaneous drift v and instantaneous variance s of the short-term interest rate process as well as the market price of risk q. Since v and s are parameters of the observable short-term interest rate process, they can be obtained by statistical analysis of market data. As previously mentioned, the market risk premium is not directly observable. Although it may be calculated from equation (1.12) as $q(r,t) = (\mu(t,T) - r(t))/\sigma(t,T)$, a more direct method of estimation may be applied. Once v and s are determined, the market risk premium may be estimated from the slope of the yield curve at the origin. From equation (1.18) we have P as a function of T and $z(T)$ and hence:

$$\begin{aligned} dP &= P_T\,dT + P_z\,dz + \tfrac{1}{2}P_{zz}\,dzdz \\ &= (-r - \tfrac{1}{2}\,q^2)P\,dT + qP\,dz + \tfrac{1}{2}\,q^2P\,dT \\ &= -rP\,dT + qP\,dz \end{aligned}$$

Therefore $\frac{\partial P}{\partial T} = -rP$ and $\frac{\partial^2 P}{\partial T^2} = -\frac{\partial(rP)}{\partial T}$. Also:

$$\begin{aligned} d(rP) &= r\,dP + P\,dr + drdP \\ &= r(-rP\,dT + qP\,dz) + P(v\,dT + s\,dz) + sqP\,dzdz \\ &= (-r^2 + v + sq)P\,dT + (rq+s)P\,dz \\ \Longrightarrow \frac{\partial(rP)}{\partial T} &= (-r^2 + v + sq)P \end{aligned}$$

and hence

$$\frac{\partial^2 P}{\partial T^2} = \big(r^2(T) - v(r(T),T) - s(r(T),T)\,q(r(T),T)\,\big)\;P$$

so

$$\left.\frac{\partial^2 P}{\partial T^2}\right|_{T=t} = r^2(t) - v(r(t),t) - s(r(t),t)\,q(r(t),t) \tag{1.19}$$

However, from equation (1.1) we have:

$$\begin{aligned}
R(t,\tau) &= -\tfrac{1}{\tau}\log P(t,T) \\
P(t,T) &= e^{-\tau R(t,\tau)} \\
\frac{\partial^2 P}{\partial T^2} &= -\left(2\frac{\partial R}{\partial \tau} + \tau\frac{\partial^2 R}{\partial \tau^2}\right) e^{-\tau R(t,\tau)} + \left(-\tau\frac{\partial R}{\partial \tau} - R(t,\tau)\right)^2 e^{-\tau R(t,\tau)} \\
\left.\frac{\partial^2 P}{\partial T^2}\right|_{T=t} &= \left(-2\left.\frac{\partial R}{\partial \tau}\right|_{\tau=0} + R^2(t,\tau)\Big|_{\tau=0}\right) e^{-\tau R(t,\tau)}\Big|_{\tau=0} \\
&= r^2(t) - 2\left.\frac{\partial R}{\partial \tau}\right|_{\tau=0} \qquad \text{by (1.3)}
\end{aligned}$$

Equating to (1.19) we have:

$$2\left.\frac{\partial R}{\partial \tau}\right|_{\tau=0} = v(r(t),t) + s(r(t),t)q(r(t),t)$$

$$\Rightarrow q(r(t),t) = \frac{2\left.\frac{\partial R}{\partial \tau}\right|_{\tau=0} - v(r(t),t)}{s(r(t),t)}$$

and hence we have found the market price of risk in terms of the slope of the term structure at the origin.

1.5. Specific model

Vasicek specifies the required parameters for the short-term interest rate process and market price of risk, to derive an explicit term structure of interest rates. He assumes the market price of risk is constant:

$$q(r,t) = q$$

and that the short-term interest rate follows an Ornstein–Uhlenbeck process:

$$dr = \alpha(\gamma - r)\,dt + s\,dz \tag{1.20}$$

where α, γ and s are constants with $\alpha > 0$. This is often referred to as an elastic random walk which is a Markovian process with normally distributed increments. The instantaneous drift $\alpha(\gamma - r)$, displays mean reversion, with the short-term interest rate being pulled to its long-term mean, γ, with magnitude proportional to its deviation from this long-term mean. This implies that the Ornstein–Uhlenbeck process is characterised by a stationary distribution, unlike a random walk (Wiener process) which is not stable and can diverge to infinite values[4].

[4] Vasicek emphasises that he is not trying to provide the best characterisation of the short-term interest rate process, but is merely specifying an example in the absence of conclusive empirical results.

Substituting (1.20) into the bond price PDE (1.13), the term structure equation becomes:

$$\frac{\partial P}{\partial t} + (\alpha(\gamma - r) + sq)\frac{\partial P}{\partial r} + \tfrac{1}{2}s^2\frac{\partial^2 P}{\partial r^2} - rP = 0, \quad t \le T \tag{1.21}$$

This is a linear, second-order partial differential equation, which can easily be solved by applying the boundary condition (1.14). Allowing the bond price to be of the form:

$$P(r(t), t, T) = A(t,T)e^{-B(t,T)r(t)} \tag{1.22}$$

equation (1.21) becomes[5]:

$$A_t e^{-Br} - rAB_t e^{-Br} - (\alpha(\gamma - r) + sq)ABe^{-Br} + \tfrac{1}{2}s^2AB^2e^{-Br} - rAe^{-Br} = 0$$

$$\begin{aligned} \Rightarrow A_t - (\alpha\gamma + sq)AB + \tfrac{1}{2}s^2AB^2 &= rA + rAB_t - \alpha rAB \\ &= rA(1 + B_t - \alpha B) \end{aligned}$$

Since the right-hand side is a function of the short-term interest rate r, while the left-hand side is a function of t and T only, the following must hold:

$$1 + B_t - \alpha B = 0 \tag{1.23}$$

$$A_t - (\alpha\gamma + sq)AB + \tfrac{1}{2}s^2AB^2 = 0 \tag{1.24}$$

Now, solving (1.23) with boundary condition $B(T,T) = 0$ gives:

$$B(t,T) = \frac{1}{\alpha}(1 - e^{-\alpha(T-t)}) \tag{1.25}$$

Rearranging (1.24), we have:

$$\begin{aligned} A_t + AB\left(\tfrac{1}{2}s^2B - \alpha\gamma - sq\right) &= 0 \\ \frac{dA}{A} + B\left(\tfrac{1}{2}s^2B - \alpha\gamma - sq\right)dt &= 0 \\ \int_t^T \frac{dA}{A} + \int_t^T \left(\tfrac{1}{2}s^2B^2(\mu,T) - (\alpha\gamma + sq)B(\mu,T)\right)d\mu &= 0 \\ \Rightarrow \ln A(T,T) - \ln A(t,T) & \\ + \frac{1}{2}\frac{s^2}{\alpha^2}\left(\mu - \frac{2}{\alpha}e^{-\alpha(T-\mu)} + \frac{1}{2\alpha}e^{-2\alpha(T-\mu)}\right)\Bigg|_{\mu=t}^{\mu=T} & \\ - \left(\gamma + \frac{sq}{\alpha}\right)\left(\mu - \frac{1}{\alpha}e^{-\alpha(T-\mu)}\right)\Bigg|_{\mu=t}^{\mu=T} &= 0 \end{aligned}$$

and hence

$$\begin{aligned} \ln A(t,T) = {} & \frac{s^2}{2\alpha^2}\left(T - t - \frac{2}{\alpha}(1 - e^{-\alpha(T-t)}) + \frac{1}{2\alpha}(1 - e^{-2\alpha(T-t)})\right) \\ & - \left(\gamma + \frac{sq}{\alpha}\right)\left(T - t - \frac{1}{\alpha}(1 - e^{-\alpha(T-t)})\right) \end{aligned}$$

[5]Here A_t and B_t denote the derivatives with respect to t.

$$
\begin{aligned}
(1.26) \qquad = \quad & \left(\gamma + \frac{sq}{\alpha} - \frac{s^2}{2\alpha^2}\right)\left(\frac{1}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) - (T-t)\right) \\
& -\frac{s^2}{4\alpha^3}\left(1 - e^{-\alpha(T-t)}\right)^2
\end{aligned}
$$

Substituting (1.25) and (1.26) into (1.22), to obtain the bond price:

$$
\begin{aligned}
P(r,t,T) = \exp\Bigg[& \left(\gamma + \frac{sq}{\alpha} - \frac{s^2}{2\alpha^2}\right)\left(\frac{1}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) - (T-t)\right) \\
& -\frac{s^2}{4\alpha^3}\left(1 - e^{-\alpha(T-t)}\right)^2 - \frac{r}{a}\left(1 - e^{-\alpha(T-t)}\right)\Bigg]
\end{aligned}
$$

Equations (1.7), (1.8) and (1.9) give the dynamics of the bond price process and the mean and standard deviation of the instantaneous rate of return of a time T maturity bond. Substitute the above bond price formula to calculate the mean and standard deviation in terms of the short-term interest rate parameters and market price of risk:

$$
\begin{aligned}
\mu(t,T) &= r(t) + \frac{sq}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) = r(t) + sqB \\
\sigma(t,T) &= \frac{s}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) = sB \quad \forall\, t \le T
\end{aligned}
$$

One can see that the standard deviation (and hence the variance) of the instantaneous rate of return increases with the bond's term to maturity and the return in excess of the short-term interest rate is proportional to this standard deviation, with the proportionality constant being the market price of risk.

Now define $R(\infty) = \gamma + \frac{sq}{\alpha} - \frac{s^2}{2\alpha^2}$ to be the long-term rate of interest. The bond price becomes:

$$
\begin{aligned}
(1.27) \quad P(r,t,T) = \exp\Bigg[& \frac{1}{\alpha}\left(1 - e^{-\alpha(T-t)}\right)(R(\infty) - r) - (T-t)\,R(\infty) \\
& -\frac{s^2}{4\alpha^3}\left(1 - e^{-\alpha(T-t)}\right)^2\Bigg]
\end{aligned}
$$

and from equation (1.1) the form of the term structure is:

$$
\begin{aligned}
(1.28) \qquad R(t,\tau) = \quad & R(\infty) + (r(t) - R(\infty))\,\frac{1}{\alpha\tau}\left(1 - e^{-\alpha\tau}\right) \\
& + \frac{s^2}{4\alpha^3\tau}\left(1 - e^{-\alpha\tau}\right)^2, \quad \tau \ge 0
\end{aligned}
$$

Since (1.28) reduces to the short-term interest rate $r(t)$ for $\tau = 0$ and $\lim_{\tau\to\infty} R(t,\tau) = R(\infty)$ it is consistent with previous definitions. Clearly the attainable shapes of the term structure depend on the value of $r(t)$. See Figure 1.1. Let us examine the how the value of $r(t)$ affects the slope of the term structure. From (1.28):

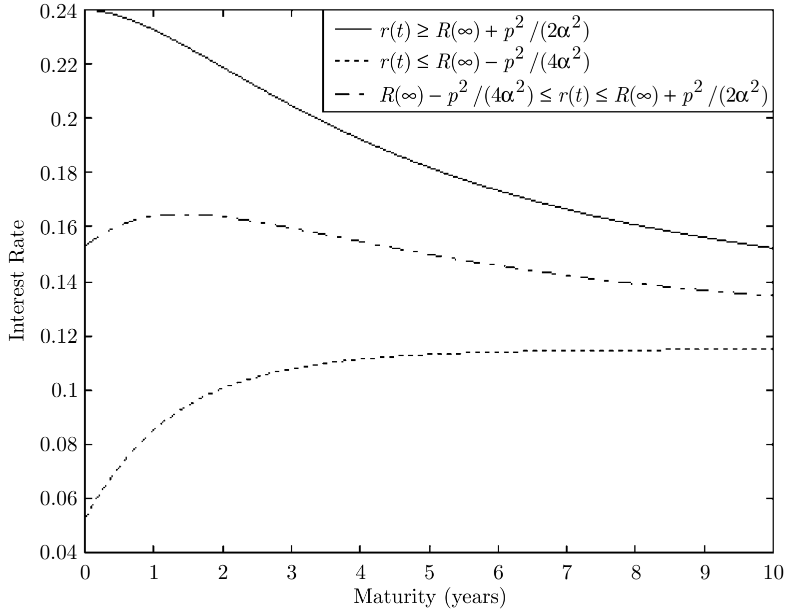

FIGURE 1.1. Possible shapes of the term structure. $\gamma = 0.14$, $\alpha = 0.5$, $s = 0.25$, $q = 0.2$

$$
\frac{\partial R}{\partial \tau} = -\frac{(r(t) - R(\infty))}{\alpha\tau^2}\left(1 - e^{-\alpha\tau}\right) + \frac{(r(t) - R(\infty))}{\alpha\tau}\,\alpha e^{-\alpha\tau}
- \frac{s^2(1 - e^{-\alpha\tau})^2}{4\alpha^3\tau^2} + \frac{s^2(1 - e^{-\alpha\tau})}{2\alpha^3\tau}\,\alpha e^{-\alpha\tau} \tag{1.29}
$$

For a monotonically increasing term structure, we require $\frac{\partial R}{\partial \tau} \geq 0$, for all τ. From (1.29) we have:

$$
\begin{aligned}
\frac{\partial R}{\partial \tau} &= \frac{1}{\alpha\tau}\left(-\frac{(1 - e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right)\right]\right. \\
&\quad \left. + e^{-\alpha\tau}\left[\alpha\left(r(t) - R(\infty)\right) + \frac{s^2}{2\alpha}\left(1 - e^{-\alpha\tau}\right)\right]\right) \\
&\geq \frac{1}{\alpha\tau}\left(-\frac{(1 - e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right)\right]\right. \\
&\quad \left. + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right)\right]\right) \\
&= \frac{1}{\alpha\tau}\left[\alpha e^{-\alpha\tau} - \frac{(1 - e^{-\alpha\tau})}{\tau}\right]\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right)\right]
\end{aligned}
$$

which is positive if $r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right) \leq 0$ because

$$\alpha e^{-\alpha\tau} - \frac{(1 - e^{-\alpha\tau})}{\tau} = \frac{(1 + \alpha\tau)e^{-\alpha\tau} - 1}{\tau} < 0 \quad \forall\tau$$

Hence, the term structure is monotonically increasing if:

$$\begin{aligned} & r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right) \leq 0 \\ \Leftrightarrow \qquad & r(t) \leq R(\infty) - \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right) \quad \forall\tau \end{aligned}$$

Additionally, since $\tau \in [0, \infty)$, $e^{-\alpha\tau} \in (0, 1]$ and $1 - e^{-\alpha\tau} \in [0, 1)$, so:

$$R(\infty) - \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right) \in \left(R(\infty) - \frac{s^2}{4\alpha^2}; R(\infty)\right]$$

and hence, for a monotonically increasing term structure, we require:

$$\begin{aligned} r(t) &\leq \min\left\{R(\infty) - \frac{s^2}{4\alpha^2}; R(\infty)\right\} \\ \Rightarrow r(t) &\leq R(\infty) - \frac{s^2}{4\alpha^2} \end{aligned}$$

For a monotonically decreasing term structure, we require $\frac{\partial R}{\partial \tau} \leq 0$ for all τ. Making a substitution similar to that used when determining a monotonically increasing term structure[6], we need to show:

$$\begin{aligned} \frac{\partial R}{\partial \tau} = \frac{1}{\alpha\tau}\Bigg(& -\frac{(1 - e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1 - e^{-\alpha\tau}\right)\right] \\ & + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) + \frac{s^2}{2\alpha^2}\left(1 - e^{-\alpha\tau}\right)\right]\Bigg) \\ \leq \frac{1}{\alpha\tau}\Bigg(& -\frac{(1 - e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) - \frac{s^2}{2\alpha^2}(1 - e^{-\alpha\tau})\right] \\ & + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) - \frac{s^2}{2\alpha^2}(1 - e^{-\alpha\tau})\right]\Bigg) \end{aligned}$$

Since $1 - e^{-\alpha\tau} \geq 0$ this is true if and only if:

$$\begin{aligned} & \alpha e^{-\alpha\tau}\frac{s^2}{\alpha^2} \leq \frac{(1 - e^{-\alpha\tau})}{\tau}\frac{3s^2}{4\alpha^2} \\ \Leftrightarrow \quad & 4\tau\alpha e^{-\alpha\tau} - 3 + 3e^{-\alpha\tau} \leq 0 \end{aligned} \tag{1.30}$$

[6]Here, substitute $-\frac{s^2}{2\alpha^2}(1 - e^{-\alpha T})$ for $\frac{s^2}{4\alpha^2}(1 - e^{-\alpha T})$ and $\frac{s^2}{2\alpha^2}(1 - e^{-\alpha T})$ in equation (1.29).

However, this inequality does not hold because[7]:

$$4\tau\alpha e^{-\alpha\tau} - 3 + 3e^{-\alpha\tau} \geq 0$$

for some interval $\tau \in [0, \tau^*]$ where τ^* depends on the value of the speed of reversion, α. Hence the above substitution is inconsistent with our argument. However, substituting $-\frac{s^2}{2\alpha^2}$ for $\frac{s^2}{4\alpha^2}(1-e^{-\alpha\tau})$ and $\frac{s^2}{2\alpha^2}(1-e^{-\alpha\tau})$ in equation (1.29) we require:

$$\begin{aligned}\frac{\partial R}{\partial \tau} &= \frac{1}{\alpha\tau}\Bigg(-\frac{(1-e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}\left(1-e^{-\alpha\tau}\right)\right] \\ &\qquad + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) + \frac{s^2}{2\alpha^2}\left(1-e^{-\alpha\tau}\right)\right]\Bigg) \\ &\leq \frac{1}{\alpha\tau}\Bigg(-\frac{(1-e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) - \frac{s^2}{2\alpha^2}\right] \\ &\qquad + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) - \frac{s^2}{2\alpha^2}\right]\Bigg)\end{aligned}$$

This inequality holds if and only if:

$$\begin{aligned} \alpha e^{-\alpha\tau}\frac{s^2}{2\alpha^2}(2-e^{-\alpha\tau}) \;&\leq\; \frac{(1-e^{-\alpha\tau})}{\tau}\frac{s^2}{4\alpha^2}(3-e^{-\alpha\tau}) \\ \iff\; 4\tau\alpha e^{-\alpha\tau} - 2\tau\alpha e^{-2\alpha\tau} - 3 + 4e^{-\alpha\tau} - e^{-2\alpha\tau} \;&\leq\; 0 \end{aligned} \tag{1.31}$$

This inequality is examined in the appendix and can be shown to hold[8] for all values of τ. Hence we can write the following:

$$\begin{aligned}\frac{\partial R}{\partial \tau} &= \frac{1}{\alpha\tau}\left(-\frac{(1-e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) + \frac{s^2}{4\alpha^2}(1-e^{-\alpha\tau})\right]\right. \\ &\qquad \left. + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) + \frac{s^2}{2\alpha^2}(1-e^{-\alpha\tau})\right]\right) \\ &\leq \frac{1}{\alpha\tau}\left(-\frac{(1-e^{-\alpha\tau})}{\tau}\left[r(t) - R(\infty) - \frac{s^2}{2\alpha^2}\right]\right. \\ &\qquad \left. + \alpha e^{-\alpha\tau}\left[r(t) - R(\infty) - \frac{s^2}{2\alpha^2}\right]\right) \\ &= \frac{1}{\alpha\tau}\left(\alpha e^{-\alpha\tau} - \frac{(1-e^{-\alpha\tau})}{\tau}\right)\left(r(t) - R(\infty) - \frac{s^2}{2\alpha^2}\right) \\ &\leq 0\end{aligned}$$

[7]This result is shown numerically in the appendix.

[8]This statement is true only if $\alpha > 0$. However this is an assumption in the formulation of the model in (1.20).

if $r(t) - R(\infty) - \frac{s^2}{2\alpha^2} \geq 0$ since $\alpha e^{-\alpha\tau} - \frac{(1-e^{-\alpha\tau})}{\tau} \leq 0$. Hence the slope is monotonically decreasing if:

$$r(t) - R(\infty) - \frac{s^2}{2\alpha^2} \geq 0$$

that is: $$r(t) \ \geq \ R(\infty) + \frac{s^2}{2\alpha^2}$$

The yield curve takes on a humped shape for values of $r(t)$ such that:

$$R(\infty) - \frac{s^2}{4\alpha^2} \ \leq \ r(t) \ \leq \ R(\infty) + \frac{s^2}{2\alpha^2}$$

1.6. Conclusion

Few of the early studies of asset prices within an equilibrium economy setting were applicable to interest rates, mainly focusing on stock prices. If we accept Vasicek's implicit assumption that the functional form of the short-term interest rate process and market price of risk are in fact consistent with an economic equilibrium, then his work may be seen as a complete characterisation of the interest rate term structure in such an economy. This simple model has been praised for its incorporation of reversion to a long-run mean and the ability to produce an analytic representation of discount bond prices. On the other hand, it has been criticised for allowing interest rates to become negative and not providing a mechanism by which the initial, market-observed term structure may be reproduced.

Appendix

We need to determine if inequality (1.30) holds. That is if:

$$4\tau\alpha e^{-\alpha\tau} - 3 + 3e^{-\alpha\tau} \leq 0$$

with $\alpha > 0$. Using the simple Matlab code below, we may determine the curve of the above function. The exact shape of the curve depends on the value of the reversion speed α. Figure 1.2 shows the curve for $\alpha = 0.15$. The above inequality does not hold for all values of τ. The interval on which $4\tau\alpha e^{-\alpha\tau} - 3 + 3e^{-\alpha\tau} \geq 0$ depends on the value of α.

```
alpha = 0.15;
y_ans = [];
x = [];
for T = 0 : 50
   T_2 = T * 0.1;
   x = [x, T_2];
   y = 4 * T_2 * alpha * exp(-alpha * T_2)
         - 3 + 3 * exp(-alpha * T_2);
   y_ans = [y_ans, y];
end
plot(x, y_ans), xlabel('τ'), ylabel('4ταe^{-ατ} + 3e^{-ατ} − 3')
```

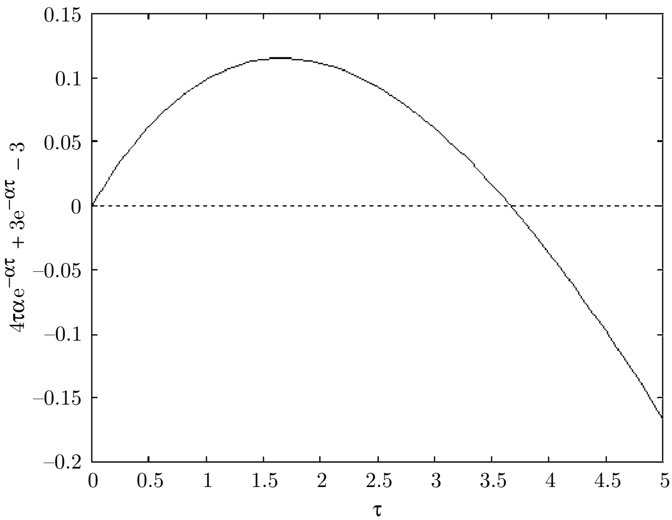

FIGURE 1.2. Shape of curve $4\tau\alpha e^{-\alpha\tau} + 3e^{-\alpha\tau} - 3$ for $\alpha = 0.15$

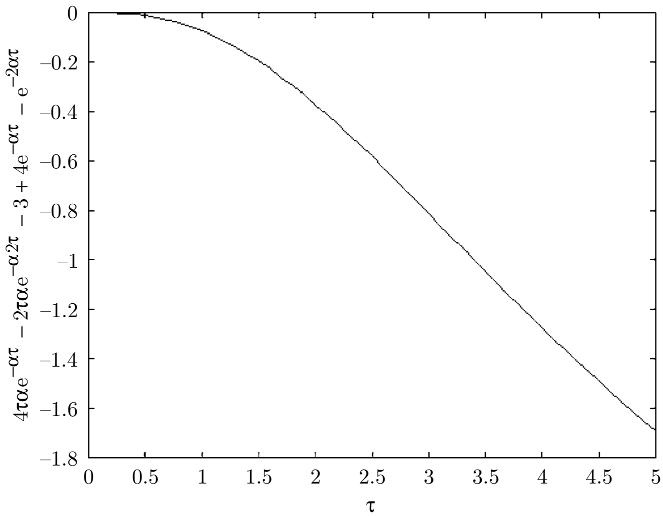

FIGURE 1.3. Shape of curve $4\tau\alpha e^{-\alpha\tau} - 2\tau\alpha e^{-2\alpha\tau} - 3 + 4e^{-\alpha\tau} - e^{-2\alpha\tau}$ for $\alpha = 0.45$

We require to show that inequality (1.31) holds. That is:

$$4\tau\alpha e^{-\alpha\tau} - 2\tau\alpha e^{-2\alpha\tau} - 3 + 4e^{-\alpha\tau} - e^{-2\alpha\tau} \;\leq\; 0$$

The Matlab code below determines the shape of this function on the interval $\tau \in [0, 5]$. Again the exact shape of the curve depends on the choice of α. We can show that the above inequality hold for all values of τ; however, α must be positive. $\alpha < 0$ implies negative reversion speed and causes a contradiction of the inequality. Figure 1.3 shows the curve for $\alpha = 0.45$.

```
alpha = 0.45;
y_ans = [];
x = [];
for T = 0 : 50
   T_2 = T * 0.1;
   x = [x, T_2];
   y = 4 * T_2 * alpha * exp(-alpha * T_2) - 2 * T_2 * alpha
         * exp(-2 * alpha * T_2) - 3 + 4 * exp(-alpha * T_2)
         - exp(-2 * alpha * T_2);
   y_ans = [y_ans, y];
end
plot(x,y_ans), xlabel('τ'), ylabel('4ταe^{−ατ} − 2ταe^{−α2τ} − 3 + 4e^{−ατ}
   −e^{−2ατ}')
```

CHAPTER 2

The Cox, Ingersoll and Ross Model

Cox, Ingersoll and Ross (CIR) view the problem of interest rate modelling as one in "general equilibrium theory" [**18**]. Anticipation of future events, risk preferences, other investment alternatives and consumption preferences all affect the term structure. CIR make use of a general equilibrium asset pricing model to endogenously determine the stochastic process followed by the short-term interest rate and the partial differential equation satisfied by the value of any contingent claim. Bond prices are then determined as solutions to this partial differential equation, contingent on the underlying short-term interest rate.

2.1. General equilibrium in a simple economy

In describing the equilibrium economy, Cox, Ingersoll and Ross [**18**] integrate real and financial markets. The specification of endogenous production and randomly changing technology allows randomly changing investment opportunities. Within a continuous time model, this characteristic allows the inclusion of effects that cannot be approximated in static, single period specifications of the economy.

Assumption 1. There is a single physical good which may be allocated to investment or for consumption. All values are expressed in terms of this good.

Assumption 2. Production possibilities are represented as a set of n linear activities. A vector of amounts η, denominated in units of the physical good, invested in production, evolves according to a stochastic process of the form:

$$d\eta(t) = I_\eta \alpha(Y,t)\,dt + I_\eta G(Y,t)\,dw(t) \tag{2.1}$$

where

$w(t) - (n+k)$-dimensional Wiener process,

$Y - k$-dimensional vector of state variables,

$I_\eta - (n \times n)$-dimensional diagonal matrix function of η. Here the i^{th} diagonal element is the i^{th} component of η,

$\alpha(Y,t)$ – n-dimensional vector of rates of return on production,

$G(Y,t)$ – $n \times (n+k)$-dimensional matrix representing the standard deviation of rates of return on production. Hence GG' is the covariance matrix of rates of return on production.

The stochastic process (2.1) describes the growth of an initial investment, given that output is continuously reinvested.

Assumption 3. The k-dimensional vector of state variables Y, evolves according to the system of stochastic equations:

$$dY(t) = \mu(Y,t)\,dt + S(Y,t)\,dw(t) \tag{2.2}$$

where

$\mu(Y,t)$ – k-dimensional vector of drifts of the state variables,

$S(Y,t)$ – $k \times (n+k)$-dimensional matrix of standard deviations of the state variables. Hence SS' is the covariance matrix of changes in state variables.

These assumptions imply that the probability distribution of current output depends on the current level of the state variables, which themselves change randomly over time. Hence the development of the state variables Y determines the future available production opportunities. Unless GS' is a null matrix[1], changes in state variables are correlated with returns on the production processes.

Assumption 4. Entry to all production processes is free.

Assumption 5. The market for instantaneous borrowing and lending exists. This occurs at a rate r which is determined as part of the equilibrium in the economy.

Assumption 6. There is a market for a variety of contingent claims to some amounts of the good. The payoffs of these claims may depend on aggregate wealth and level of the state variables. The values of the claims depend on the same variables that describe the state of the economy; hence the movement of the value of claim i, F^i, is described by the stochastic differential equation:

$$dF^i = (F^i\beta_i - \delta_i)\,dt + F^i h_i\,dw(t) \tag{2.3}$$

[1] GS' is the covariance matrix of changes in state variables and returns on production processes.

where

$F^i\beta_i - \delta_i$ – mean price drift. δ_i is the payout flow received, hence $F^i\beta_i$ is the total mean return,

h_i – $1 \times (n+k)$-dimensional vector of standard deviations, hence $h_i h_i'$ is the variance of the rate of return.

The equilibrium value of β_i, the rate of return on the contingent claim, and r, the risk-free rate of interest, are determined endogenously.

Assumption 7. There is a fixed number of identical individuals, each wishing to maximise their objective function of the form:

$$\mathbb{E}\left[\int_t^{t'} U\big(C(s), Y(s), s\big)\, ds\right] \tag{2.4}$$

where

$\mathbb{E}[\,\cdot\,]$ – expectation conditional on current wealth and state of the economy,

$C(s)$ – consumption flow at time s,

U – von Neumann–Morgenstern utility function. It is assumed to be increasing, strictly concave, twice differentiable and to satisfy:

$$\big|\, U\big(C(s), Y(s), s\big)\big| \le k_1\big(1 + C(s) + |Y(s)|\big)^{k_2}$$

for some $k_1, k_2 > 0$.

Assumption 8. Trading and investment take place continuously, at equilibrium prices only and free of transaction costs.

Consider the problem of allocating an individual's wealth to investment opportunities. If contingent claims exist, the solution will usually not be unique. We select a basis to be a set of production opportunities and contingent claims, such that any other contingent claims may be expressed as a linear combination thereof. Hence define the opportunity set as a basis of investment opportunities consisting of n production activities and k contingent claims. Individuals may allocate wealth among these $(n + k)$ basis opportunities and the $(n + k + 1)^{th}$ opportunity: riskless lending or borrowing. Since allocation to non-basis contingent claims may be replicated by a portfolio of basis claims, a unique allocation to basis investment opportunities is sufficient for valuation purposes.

Define:

W – current total wealth,

$a_i W$ – amount of wealth invested in the i^{th} production process,

$b_i W$ – amount of wealth invested in the i^{th} contingent claim.

Hence, for each individual, we wish to choose controls aW, bW and C to maximise expected lifetime utility, subject to the budget constraint denoted as[2]:

$$
\begin{aligned}
dW &= \left[\sum_{i=1}^{n} a_i W(\alpha_i - r) + \sum_{i=1}^{k} b_i W(\beta_i - r) + rW - C\right] dt \\
&\quad + \sum_{i=1}^{n} a_i W\left(\sum_{j=1}^{n+k} g_{ij}\, dw_j\right) + \sum_{i=1}^{k} b_i W\left(\sum_{j=1}^{n+k} h_{ij}\, dw_j\right) \\
&= W\mu(W)\, dt + W \sum_{j=1}^{n+k} q_j\, dw_j
\end{aligned} \tag{2.5}
$$

To maximise the utility function (2.4), the control must be a measurable function. This means the control, at any time, may only depend on information available at that time. Hence the following lemma defines the basic optimality condition for an individual's allocation (control) problem[3] (see [**17**]).

LEMMA. Let $J(W, Y, t)$ be the solution to the Bellman equation of the form:

$$
\max_{v \in V} \left[L^v(t)J + U(v, Y, t)\right] + J_t = 0 \tag{2.6}
$$

for $(t, W, Y) \in \mathcal{D} \equiv [t, t') \times (0, \infty) \times R^k$ and with boundary conditions

$$
J(0, Y, t) = \mathbb{E}_{Y,t}\left[\int_t^{t'} U\big(0, Y(s), s\big)\, ds\right] \quad \text{and} \quad J(W, Y, t') = 0
$$

Then

(1) $J(W, Y, t) \geq K(v, W, Y, t)$ for any admissible control v and initial W and Y,

(2) if $\hat{v}$ is an admissible control such that

$$
L^{\hat{v}}(t)J + U(\hat{v}, Y, t) = \max_{v \in V}\left[L^v(t)J + U(v, Y, t)\right] \qquad \forall\, (t, W, Y) \in \mathcal{D}
$$

[2] Here w_j denotes the j^{th} element of $w(t)$, i.e. the j^{th} Wiener process.

[3] Here, define the following:

$$
K(v(t), W(t), Y(t), t) = \mathbb{E}\left[\int_t^{t'} U\big(v(s), Y(s), s\big)\, ds\right]
$$

where $v(s)$ is some admissible control. Also let $L^v(t)K$ be the associated differential operator defined as:

$$
\begin{aligned}
L^v(t)K &= \mu(W) W K_W + \sum_{i=1}^{k} \mu_i K_{Y_i} + \tfrac{1}{2}\, W^2 K_{WW} \sum_{i=1}^{n+k} q_i^2 \\
&\quad + \sum_{i=1}^{k} W K_{WY_i} \sum_{j=1}^{n+k} q_j s_{ij} + \tfrac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} K_{Y_i Y_i} \sum_{m=1}^{n+k} s_{im} s_{jm}
\end{aligned}
$$

then

$$J(W, Y, t) = K(\hat{v}, W, Y, t) \qquad\qquad \forall\,(t, W, Y) \in \mathcal{D}$$

and $\hat{v}$ is optimal.

Here J is called the indirect utility function.

Assumption 10. There exists a unique function J and control $\hat{v}$ satisfying the Bellman equation (2.6) and associated technical conditions.

Since investment proportions a_i and consumption C must be non-negative, necessary and sufficient conditions for maximising $\psi = L^v J + U$ as a function of C, a and b may be expressed as [**17**]:

$$\psi_C = U_C - J_W \leq 0, \tag{2.7a}$$

$$C\psi_C = 0, \tag{2.7b}$$

$$\psi_a = (\alpha - r1)WJ_W + (GG'a + GH'b)W^2 J_{WW} + GS'WJ_{WY} \leq 0, \tag{2.7c}$$

$$a'\psi_a = 0, \tag{2.7d}$$

$$\psi_b = (\beta - r1)WJ_W + (HG'a + HH'b)W^2 J_{WW} + HS'WJ_{WY} = 0 \tag{2.7e}$$

Using (2.7) and the Bellman equation (2.6) $\hat{C}$, $\hat{a}$ and $\hat{b}$ may be found in terms of W, Y and t only. $\hat{C}$, $\hat{a}$ and $\hat{b}$ are chosen, taking r, α and β as given. The set of stochastic processes $(r, \beta; a, C)$ (that is, the risk-free rate of interest, expected returns on contingent claims, production plan and consumption plan) define the equilibrium economy under the conditions $\sum a_i = 1$ and $b_i = 0$ for all i. These conditions imply that in equilibrium, the interest rate and expected rates of return on the contingent claims are such that all wealth is invested in physical production processes only [**18**].

2.2. Equilibrium risk-free rate of interest

We need to determine the equilibrium values of a, r and β. The solution to an individual's planning problem consists of an optimal physical investment policy a^*, optimal consumption plan C^* and the associated indirect utility function J^*. By writing the portfolio allocation (physical investment) part as a quadratic programming problem, CIR [**17**] determine the equilibrium interest rate to be of the form:

$$\begin{aligned} r(W, Y, t) &= a^{*\prime}\alpha + a^{*\prime}GG'a^*W\left(\frac{J_{WW}}{J_W}\right) + a^*GS'\left(\frac{J_{WY}}{J_W}\right) \\ &= a^{*\prime}\alpha - \left(\frac{-J_{WW}}{J_W}\right)\left(\frac{\operatorname{var}(W)}{W}\right) - \sum_{i=1}^{k}\left(\frac{-J_{WY_i}}{J_W}\right)\left(\frac{\operatorname{covar}(W, Y_i)}{W}\right) \end{aligned} \tag{2.8}$$

2.3. Equilibrium expected return on any contingent claim

Considering optimal physical investment a^*, and substituting the equilibrium interest rate (2.8) into (2.7e) we get the equilibrium expected return on any contingent claim:

$$\beta(W,Y,t) = \big(a^{*\prime}\alpha\big)1 + \left(\frac{1}{J_W}\right)\big[\big(a^{*\prime}GS'J_{WY}\big)1 - HS'J_{WY}\big] + \left(\frac{WJ_{WW}}{J_W}\right)\big[\big(a^{*\prime}GG'a^*\big)1 - HG'a^*\big] \tag{2.9}$$

Now, applying Ito's lemma to $F(W,Y,t)$, and making use of (2.5) and (2.2):

$$\begin{aligned} dF &= \frac{\partial F}{\partial t}\,dt + \frac{\partial F}{\partial W}\,dW + \tfrac{1}{2}\frac{\partial^2 F}{\partial W^2}\,dW dW + \sum_{j=1}^{k}\frac{\partial F}{\partial Y_j}\,dY_j \\ &\quad + \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{\partial^2 F}{\partial Y_i\,\partial Y_j}\,dY_i dY_j + \sum_{j=1}^{k}\frac{\partial^2 F}{\partial W\,\partial Y_j}\,dW dY_j \\ &= \mu_F\,dt + \left(F_W\,aWG + \sum_{j=1}^{k}F_{Y_j}\,S\right)dw(t) \end{aligned} \tag{2.10}$$

with:

$$\begin{aligned} \mu_F = F_t + F_W\big(a'\alpha W - C\big) + \sum_{j=1}^{k}F_{Y_j}\mu_j(Y,t) + \tfrac{1}{2}F_{WW}\,\mathrm{var}\,(W) \\ + \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}F_{Y_iY_j}\,\mathrm{covar}\,(Y_i,Y_j) + \sum_{j=1}^{k}F_{WY_j}\,\mathrm{covar}\,(W,Y_j) \end{aligned}$$

Now comparing the volatility components of (2.10) and (2.3) we have:

$$FH = F_W\,aWG + \sum_{j=1}^{k}F_{Y_j}\,S$$

Substitution into (2.9) gives:

$$\begin{aligned} \beta F = \big(a^{*\prime}\alpha\big)F + a^{*\prime}GG'a^*W\left(\frac{J_{WW}}{J_W}\right)F + a^{*\prime}GS'\left(\frac{J_{WY}}{J_W}\right)F \\ + F_W\left[\left(\frac{-J_{WW}}{J_W}\right)\mathrm{var}\,(W) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\mathrm{covar}\,(W,Y_j)\right] \\ + \sum_{i=1}^{k}F_{Y_i}\left[\left(\frac{-J_{WW}}{J_W}\right)\mathrm{covar}\,(W,Y_i) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\mathrm{covar}\,(Y_i,Y_j)\right] \end{aligned}$$

which by (2.8) becomes:

(2.11)

$$\beta F = rF + F_W\left[\left(\frac{-J_{WW}}{J_W}\right)\text{var}\,(W) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\text{covar}\,(W,Y_j)\right]$$
$$+\sum_{i=1}^{k}F_{Y_i}\left[\left(\frac{-J_{WW}}{J_W}\right)\text{covar}\,(W,Y_i) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\text{covar}\,(Y_i,Y_j)\right]$$

So the equilibrium expected return on any contingent claim may be written as the risk-free return rF, plus a linear combination of the first derivatives of the contingent claim price with respect to wealth W, and the state variables Y. The coefficients of these derivatives are independent of the contractual specification for that claim; hence they are the same for all contingent claims. CIR [**17**] explain that these coefficients may be interpreted as factor risk premia[4].

2.4. Value of any contingent claim in equilibrium

Now comparing the drift coefficients in (2.10) and (2.3) we have:

$$\beta F - \delta = \mu_F = F_t + F_W\big(a'\alpha W - C\big) + \sum_{j=1}^{k}F_{Y_j}\mu_j(Y,t) + \tfrac{1}{2}F_{WW}\,\text{var}\,(W)$$
$$+\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}F_{Y_iY_j}\,\text{covar}\,(Y_i,Y_j) + \sum_{j=1}^{k}F_{WY_j}\,\text{covar}\,(W,Y_j)$$

Making use of (2.11) yields:

$$rF + F_W\left[\left(\frac{-J_{WW}}{J_W}\right)\text{var}\,(W) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\text{covar}\,(W,Y_j)\right]$$
$$+\sum_{i=1}^{k}F_{Y_i}\left[\left(\frac{-J_{WW}}{J_W}\right)\text{covar}\,(W,Y_i) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\text{covar}\,(Y_i,Y_j)\right] - \delta$$
$$= F_t + F_W\big(a'\alpha W - C\big) + \sum_{j=1}^{k}F_{Y_j}\mu_j(Y,t) + \tfrac{1}{2}F_{WW}\,\text{var}\,(W)$$
$$+\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}F_{Y_iY_j}\,\text{covar}\,(Y_i,Y_j) + \sum_{j=1}^{k}F_{WY_j}\,\text{covar}\,(W,Y_j)$$

[4]Specifically from (2.11), the risk premium for the i^{th} state variable Y_i is

$$\left[\left(\frac{-J_{WW}}{J_W}\right)\text{covar}\,(W,Y_i) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\text{covar}\,(Y_i,Y_j)\right]$$

Rearranging terms we have a partial differential equation for the price of any contingent claim[5]:

$$\begin{aligned}\tfrac{1}{2}F_{WW}\operatorname{var}(W) + \sum_{j=1}^{k} F_{WY_j}\operatorname{covar}(W,Y_j) + \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k} F_{Y_iY_j}\operatorname{covar}(Y_i,Y_j)\\ + \sum_{i=1}^{k} F_{Y_i}\left[\mu_i(Y,t) - \left(\frac{-J_{WW}}{J_W}\right)\operatorname{covar}(W,Y_i) + \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\operatorname{covar}(Y_i,Y_j)\right]\\ +\big(rW - C^*\big)F_W + F_t - rF + \delta = 0 \qquad (2.12)\end{aligned}$$

This valuation equation holds for any contingent claim. Specific terminal and boundary conditions as well as the structure of δ, the payout flow, define the unique characteristics of a claim.

2.5. A more specialised economy

For the problem of modelling the interest rate term structure, CIR specialise the economy. They restrict the class of utility functions to those having constant relative risk aversion[6]. Specifically, the utility function is required to be logarithmic and independent of the state variable Y, hence:

$$U\big[C(s), s\big] = e^{-\rho s}\ln C(s) \qquad (2.13)$$

where ρ is a constant discount factor.

CIR [**18**] show that for this specialised case the indirect utility function takes the form:

$$J(W,Y,t) = h(t)U(W,t) + g(Y,t)$$

for some functions $h(t)$ and $g(Y,t)$. By the results of the earlier Lemma

$$h(t) = \frac{1 - e^{(-\rho(t'-t))}}{\rho}$$

Hence we have:

$$\frac{-WJ_{WW}}{J_W} = 1 \quad \text{and} \quad \frac{-J_{WY}}{J_W} = 0$$

Substituting into (2.8), the form of the equilibrium interest rate reduces to:

$$r = a^{*\prime}\alpha - a^{*\prime}GG'a^* \qquad (2.14)$$

[5]First group the coefficients of F_W and make use of (2.8) to give:

$$\left[a'\alpha W - \left(\frac{-J_{WW}}{J_W}\right)\operatorname{var}(W) - \sum_{j=1}^{k}\left(\frac{-J_{WY_j}}{J_W}\right)\operatorname{covar}(W,Y_j) - C\right]F_W = \big(rW - C^*\big)F_W$$

[6]That is neither the interest rate, nor the security risk premia depend on the level of investor wealth.

Similarly, by (2.11), the return on any contingent claim simplifies to:

$$\beta F = rF + F_W a^{*\prime} G G' a^* W + \sum_{i=1}^{k} F_{Y_i} a^{*\prime} G S' \tag{2.15}$$

For the purposes of developing a model of the interest rate term structure, CIR assume that the contractual terms of all securities are free of explicit dependence on wealth. This implies that the partial derivatives, with respect to wealth, of all securities, equal zero (i.e. $F_W = F_{WW} = F_{WY} = 0$). As shown above the risk-free rate and factor risk premia[7] are also independent of wealth. Due to this additional restriction the valuation equation for contingent claims (2.12) reduces to:

$$\begin{aligned} \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k} F_{Y_i Y_j} \operatorname{covar}(Y_i, Y_j) + \sum_{i=1}^{k} F_{Y_i}\left[\mu_i(Y,t) - a^{*\prime} G S'\right] & \\ + F_t - rF + \delta = 0 & \end{aligned} \tag{2.16}$$

2.6. Term structure model

To model the term structure CIR use equation (2.16) and make several simplifying assumptions about technological change in their specialised equilibrium economy.

Assumption 1. The change in production opportunities over time is determined by a single state variable Y.

Assumption 2. The means and variances of the rates of return on the production processes are proportional to Y. Hence the state variable Y determines the rate of evolution of capital and neither the means nor variances will dominate the portfolio decisions for increasing values of Y.

Assumption 3. The state variable Y follows the following stochastic process[8]:

$$dY(t) = [\xi Y + \zeta]\, dt + \nu\sqrt{Y} dw(t)$$

where ξ and ζ constants, $\zeta \geq 0$ and ν a vector of constants.

Incorporating the above assumptions into their economic model, CIR [**18**] derive an explicit formula for the equilibrium interest rate in terms of the state variable Y, the parameters of its stochastic process and the means and variances of the rates of return on the production processes in the economy. Calculating the drift and variance of this equilibrium interest rate and defining a new Wiener process $z(t)$, such that:

$$\sigma\sqrt{r} dz(t) \equiv \nu\sqrt{Y} dw(t)$$

[7] In (2.15) the factor risk premia, that is the coefficients of F_{Y_i}, reduce to $a^{*\prime} GS'$.
[8] This is a specialisation of (2.2)

they specify the dynamics of the interest rate as[9]:

$$dr = \kappa(\theta - r)dt + \sigma\sqrt{r}dz(t) \tag{2.17}$$

where $\kappa, \theta \geq 0$. (2.17) represents a continuous time first-order autoregressive process where the stochastic interest rate is pulled to its long-term mean θ, with speed κ. Imposing an additional constraint that $2\kappa\theta \geq \sigma^2$ (see §2.7.6) ensures that the rate of interest cannot become negative. The interest rate structure implied by (2.17) displays the following characteristics:

- negative interest rates are prevented,
- a zero rate of interest can become positive again,
- the level of absolute variance increases with increasing interest rates,
- the interest rate displays a steady state distribution.

Within the same economic framework CIR determine the factor risk premium in terms of the above-mentioned economic variables. Together with all the simplifying assumptions, the factor risk premium λ, is substituted into (2.16) to give the fundamental bond equation which, in equilibrium, must be satisfied by any zero coupon bond:

$$\tfrac{1}{2}\sigma^2 r P_{rr} + \kappa(\theta - r)P_r + P_t - \lambda r P_r - rP = 0 \tag{2.18}$$

with $P(r, T, T) = 1$ as the boundary condition. Here the bond price depends on one underlying stochastic variable, the short-term interest rate r, which represents the uncertainty in the underlying economy. This model proposes that the short-term interest rate is the predominant variable in determining the whole term structure. This can only be true if all the simplifying assumptions are satisfied[10].

2.7. Distribution of the interest rate

2.7.1. Mean. To calculate the conditional mean and variance of the interest rate under the CIR model[11], consider the integral form of (2.17):

$$r(s) = r(t) + \kappa \int_t^s \big(\theta - r(u)\big)\, du + \sigma \int_t^s \sqrt{r(u)}\, dz(u)$$

since $r(t)$, the current rate of interest is known. The mean of a Wiener process is zero, and so taking expectations:

$$\mathbb{E}\left[r(s)|r(t)\right] = r(t) + \kappa \int_t^s \big(\theta - \mathbb{E}\left[r(u)|r(t)\right]\big)\, du$$

[9]See [**18**] for the details of this.

[10]These assumptions can be summarised as: investors have constant relative risk aversion, uncertainty within the economy is modelled by a single variable, and the interest rate is a monotonic function of this variable.

[11]See [**49**] for detailed calculations of the mean and variance.

In differential form this equation is represented as:

$$\frac{\partial}{\partial s}\,\mathbb{E}\left[r(s)|r(t)\right] = \kappa\left(\theta - \mathbb{E}\left[r(s)|r(t)\right]\right)$$

which is a separable differential equation which may be solved as follows:

$$\begin{aligned}
\frac{d\,\mathbb{E}\left[r(s)|r(t)\right]}{\theta - \mathbb{E}\left[r(s)|r(t)\right]} &= \kappa\,ds \\
\int_t^s \frac{d\,\mathbb{E}\left[r(u)|r(t)\right]}{\theta - \mathbb{E}\left[r(u)|r(t)\right]} &= \int_t^s \kappa\,du \\
\ln\frac{\theta - \mathbb{E}\left[r(s)|r(t)\right]}{\theta - r(t)} &= -\kappa(s-t) \\
\Rightarrow \mathbb{E}\left[r(s)|r(t)\right] &= \theta + e^{-\kappa(s-t)}\big(r(t)-\theta\big) \qquad (2.19)
\end{aligned}$$

2.7.2. Variance. Calculate the stochastic differential equation satisfied by $r^2(t)$ by applying Ito's formula to (2.17). Define $f(x) = x^2$ then:

$$\begin{aligned}
df(r(t)) &= f'(r(t))dr(t) + \tfrac{1}{2}f''(r(t))dr(t)dr(t) \\
\Rightarrow d\big(r^2(t)\big) &= 2r(t)\left[\kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dz\right] \\
&\quad + \left[\kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dz\right]^2 \\
&= 2\kappa\,\theta\,r(t)dt - 2\kappa\,r^2(t)dt + 2\sigma\,r^{3/2}(t)dz + \sigma^2\,r(t)dt \\
&= (2\kappa\,\theta + \sigma^2)r(t)dt - 2\kappa\,r^2(t)dt + 2\sigma\,r^{3/2}(t)dz \\
\Rightarrow r^2(s) &= r^2(t) + (2\kappa\,\theta + \sigma^2)\int_t^s r(u)\,du - 2\kappa\int_t^s r^2(u)\,du \\
&\quad + 2\sigma\int_t^s r^{3/2}(u)\,dz(u)
\end{aligned}$$

Considering the conditional expectation of $r^2(s)$ we have:

$$\mathbb{E}\left[r^2(s)|r(t)\right] = r^2(t) + (2\kappa\,\theta + \sigma^2)\int_t^s \mathbb{E}\left[r(u)|r(t)\right]du - 2\kappa\int_t^s \mathbb{E}\left[r^2(u)|r(t)\right]du$$

Partial differentiation with respect to s yields:

$$\frac{\partial}{\partial s}\,\mathbb{E}\left[r^2(s)|r(t)\right] = (2\kappa\,\theta + \sigma^2)\,\mathbb{E}\left[r(s)|r(t)\right] - 2\kappa\,\mathbb{E}\left[r^2(s)|r(t)\right]$$

and hence:

$$\begin{aligned}
&\frac{\partial}{\partial s}\left(e^{2\kappa(s-t)}\,\mathbb{E}\left[r^2(s)|r(t)\right]\right) \\
&= 2\kappa\,e^{2\kappa(s-t)}\,\mathbb{E}\left[r^2(s)|r(t)\right] + e^{2\kappa(s-t)}\frac{\partial}{\partial s}\,\mathbb{E}\left[r^2(s)|r(t)\right] \\
&= 2\kappa\,e^{2\kappa(s-t)}\,\mathbb{E}\left[r^2(s)|r(t)\right] + 2\kappa\,\theta\,e^{2\kappa(s-t)}\,\mathbb{E}\left[r(s)|r(t)\right] \\
&\quad + \sigma^2\,e^{2\kappa(s-t)}\,\mathbb{E}\left[r(s)|r(t)\right] - 2\kappa\,e^{2\kappa(s-t)}\,\mathbb{E}\left[r^2(s)|r(t)\right] \\
&= (2\kappa\,\theta + \sigma^2)e^{2\kappa(s-t)}\,\mathbb{E}\left[r(s)|r(t)\right]
\end{aligned}$$

Integrating the above over $[t, s]$ and using (2.19) gives:

$$
\begin{aligned}
& e^{2\kappa(s-t)}\,\mathbb{E}\left[r^2(s)|r(t)\right] - r^2(t) \\
&= \int_t^s (2\kappa\,\theta + \sigma^2)\, e^{2\kappa(u-t)}\,\mathbb{E}\left[r(u)|r(t)\right]\, du \\
&= \int_t^s (2\kappa\,\theta + \sigma^2)\, e^{2\kappa(u-t)}\Big(r(t)e^{-\kappa(u-t)} + \theta\big(1 - e^{-\kappa(u-t)}\big)\Big)du \\
&= (2\kappa\,\theta + \sigma^2)r(t)\int_t^s e^{\kappa(u-t)}du + \theta\,(2\kappa\,\theta + \sigma^2)\int_t^s e^{2\kappa(u-t)} - e^{\kappa(u-t)}du \\
&= (2\kappa\,\theta + \sigma^2)(r(t) - \theta)\int_t^s e^{\kappa(u-t)}du + \theta\,(2\kappa\,\theta + \sigma^2)\int_t^s e^{2\kappa(u-t)}du \\
&= \frac{1}{\kappa}(2\kappa\,\theta + \sigma^2)(r(t) - \theta)\big(e^{\kappa(s-t)} - 1\big) + \frac{\theta}{2\kappa}(2\kappa\,\theta + \sigma^2)\big(e^{2\kappa(s-t)} - 1\big) \\
\Rightarrow\ & \mathbb{E}\left[r^2(s)|r(t)\right] \\
&= r^2(t)e^{-2\kappa(s-t)} + \frac{1}{\kappa}(2\kappa\,\theta + \sigma^2)(r(t) - \theta)e^{-\kappa(s-t)} \\
&\quad - \frac{1}{\kappa}(2\kappa\,\theta + \sigma^2)(r(t) - \theta)e^{-2\kappa(s-t)} + \frac{1}{2\kappa}\,\theta(2\kappa\,\theta + \sigma^2) \\
&\quad - \frac{1}{2\kappa}\,\theta(2\kappa\,\theta + \sigma^2)e^{-2\kappa(s-t)} \\
&= \frac{\theta\,\sigma^2}{2\kappa} + \theta^2 + (r(t) - \theta)\Big(\frac{\sigma^2}{\kappa} + 2\theta\Big)e^{-\kappa(s-t)} \\
&\quad + \big(r^2(t) + \theta^2 - 2\theta\, r(t)\big)e^{-2\kappa(s-t)} + \Big(\frac{\sigma^2}{2\kappa}\,\theta - \frac{\sigma^2}{\kappa}r(t)\Big)e^{-2\kappa(s-t)} \\
&= \frac{\theta\,\sigma^2}{2\kappa} + \theta^2 + (r(t) - \theta)\Big(\frac{\sigma^2}{\kappa} + 2\theta\Big)e^{-\kappa(s-t)} \\
&\quad + \big(r(t) - \theta\big)^2 e^{-2\kappa(s-t)} + \frac{\sigma^2}{\kappa}\Big(\frac{\theta}{2} - r(t)\Big)e^{-2\kappa(s-t)}
\end{aligned}
$$

From (2.19) we have:

$$
\big(\,\mathbb{E}\left[r(s)|r(t)\right]\big)^2 = (r(t) - \theta)^2 e^{-2\kappa(s-t)} + \theta^2 + 2\theta(r(t) - \theta)e^{-\kappa(s-t)}
$$

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

$$
\begin{aligned}
\operatorname{var}(r(s)|r(t)) &= \mathbb{E}\left[r^2(s)|r(t)\right] - \big(\,\mathbb{E}\left[r(s)|r(t)\right]\big)^2 \\
&= \frac{\theta\sigma^2}{2\kappa} + \frac{\sigma^2}{\kappa}(r(t) - \theta)e^{-\kappa(s-t)} + \frac{\sigma^2}{\kappa}\Big(\frac{\theta}{2} - r(t)\Big)e^{-2\kappa(s-t)} \qquad (2.20)
\end{aligned}
$$

2.7.3. Identifying the distribution. Define

$$r(t) = X_1^2(t) + X_2^2(t) + \cdots + X_n^2(t)$$

where each $X_i(t)$ is an Ornstein–Uhlenbeck process representing the solution to [**49**]:

$$dX_i(t) = -\tfrac{1}{2}\beta X_i(t)dt + \tfrac{1}{2}\sigma dz_i(t) \tag{2.21}$$

where $\beta > 0$ and $\sigma > 0$ are constants and z_i is a Wiener process. The conditional mean of X_i is:

$$\mathbb{E}\left[X_i(s)|X_i(t)\right] = e^{-(\beta/2)(s-t)}X_i(t) \tag{2.22}$$

The stochastic process for $r(t)$ is derived by application of Itô's formula as[12]:

$$dr(t) = \left(\frac{n\sigma^2}{4} - \beta r(t)\right)dt + \sigma\sqrt{r(t)}\, d\tilde{z}(t) \tag{2.23}$$

[12]Since $r(t) = X_1^2 + X_2^2 + \ldots + X_n^2 \equiv r(X_1, X_2, \ldots, X_n)$ we have, by the independence of the X_is:

$$r_{X_i} = 2X_i, \quad r_{X_iX_j} = \begin{cases} 2 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

where the subscripts denote partial derivatives, and so

$$\begin{aligned} dr(t) &= \sum_{i=1}^n r_{X_i} dX_i + \frac{1}{2}\sum_{i=1}^n r_{X_iX_i} dX_i dX_i \\ &= \sum_{i=1}^n 2X_i\left(-\frac{1}{2}\beta X_i(t)dt + \frac{1}{2}\sigma dz_i(t)\right) + \sum_{i=1}^n \frac{1}{4}\sigma^2 dz_i(t)dz_i(t) \\ &= \sum_{i=1}^n -\beta X_i^2 dt + \sum_{i=1}^n \sigma X_i dz_i(t) + \sum_{i=1}^n \frac{1}{4}\sigma^2 dt \\ &= -\beta r(t)dt + \sigma\sum_{i=1}^n X_i dz_i(t) + \frac{n\sigma^2}{4}dt \\ &= \left(\frac{n\sigma^2}{4} - \beta r(t)\right) dt + \sigma\sqrt{r(t)}d\tilde{z}(t) \end{aligned}$$

where $\tilde{z}(t)$ the transformed Brownian motion, is defined as:

$$\tilde{z}(t) = \sum_{i=1}^n \int_0^t \frac{X_i(t)}{\sqrt{r(t)}}\, dz_i(t)$$

To verify that $\tilde{z}(t)$ is in fact a Brownian motion, consider:

$$d\tilde{z}(t) = \sum_{i=1}^n \frac{X_i(t)}{\sqrt{r(t)}}\, dz_i(t)$$

$$\Rightarrow d\tilde{z}(t)\, d\tilde{z}(t) = \sum_{i=1}^n \frac{X_i^2(t)}{r(t)}\, dt = dt \text{ by definition of } r(t)$$

Also, since $z_i(t),\ i = 1, \ldots, n$ are martingales, then $\tilde{z}(t)$ is a martingale and we have verified that $\tilde{z}(t)$ is in fact a valid Brownian motion.

By setting $\kappa = \beta$ and $\theta = n\sigma^2/(4\beta)$ we arrive at the formulation in (2.17). Defining $r(t)$ as the sum of squares of normally distributed[13] variables $X_i,\ i = 1, \ldots, n$ leads to the correct form of its stochastic process, hence we may conclude that $r(t)$ has a non-central χ^2 distribution.

2.7.4. Chi-Square distribution. Consider a random variable Y, which may be expressed as:

$$Y = X_1^2 + X_2^2 + \ldots + X_d^2$$

where the X_is form a mutually independent set and each $X_i^2,\ i = 1, \ldots, d$ is a standard normal variable, i.e. $X_i \sim \Phi(0, 1),\ i = 1, \ldots, d$. Y has a χ^2 distribution with d degrees of freedom [**35**].

If $X_i \sim \Phi(c_i, 1),\ i = 1, \ldots, d$, that is the X_is are normally distributed with mean c_i and variance 1, then Y has a non-central χ^2 distribution with parameter of non-centrality $\varphi \equiv \sum c_i^2$ and d degrees of freedom [**35**]. The mean, variance and probability density function $f(\cdot)$, of Y are expressed as [**35**]:

$$\begin{aligned} \mathbb{E}\left[Y\right] &= d + \varphi \\ \operatorname{var}\left(Y\right) &= 2d + 4\varphi \qquad (2.24) \\ f(Y) &= \frac{e^{-\frac{1}{2}Y} e^{-\frac{1}{2}\varphi}\, Y^{\frac{1}{2}d-1}}{2^{\frac{1}{2}d}\,\Gamma(\frac{1}{2}d)} \left(1 + \frac{1}{d}\frac{Y\varphi}{2} + \frac{1}{d(d+2)}\frac{1}{2!}\left(\frac{Y\varphi}{2}\right)^2 + \cdots\right), \\ & \qquad 0 \le Y \le \infty \end{aligned}$$

where $\Gamma(\cdot)$ is the Gamma function[14].

Define the following variables:

$$\begin{aligned} c &= \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(s-t)})} \\ u &= c\, r(t)\, e^{-\kappa(s-t)} \\ \nu &= c\, r(s) \\ q &= \frac{2\kappa\,\theta}{\sigma^2} - 1 \end{aligned}$$

Now calculate the parameters required to characterise the distribution function of $r(s)$, conditional on its value at $r(t)$. From (2.23) the number of degrees of freedom are:

$$n = 4\kappa\,\theta/\sigma^2 = 2q + 2$$

[13]The X_is are normally distributed by their specification as Ornstein–Uhlenbeck processes in (2.21).

[14]The Gamma function is defined as:

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$$

and $\Gamma(\alpha) = (\alpha - 1)!$ for $\alpha \in \mathbb{N}$.

The parameter of non-centrality may be calculated as $\varphi^* = \sum c_i^2$ where $c_i,\ i = 1, \dots, n$ are the conditional means of the X_is as specified in (2.22):

$$\begin{aligned}
\varphi^* &= \sum_{i=1}^{n} \Big(\mathbb{E}\left[X_i(s) | X_i(t)\right] \Big)^2 \\
&= \sum_{i=1}^{n} e^{-\kappa(s-t)} X_i^2(t) \\
&= e^{-\kappa(s-t)} r(t)
\end{aligned}$$

From (2.19) the conditional expectation of the short-term interest rate as specified by the CIR model is:

$$\begin{aligned}
\mathbb{E}\left[r(s)|r(t)\right] &= \theta + e^{-\kappa(s-t)}\big(r(t) - \theta\big) \\
&= \theta\big(1 - e^{-\kappa(s-t)}\big) + e^{-\kappa(s-t)} r(t) \\
&= \frac{n}{2c} + \varphi^* \\
\Rightarrow \mathbb{E}\left[2c\, r(s)|r(t)\right] &= n + 2c\varphi^*
\end{aligned}$$

Similarly, the conditional variance of the short-term interest rate is specified by (2.20) as:

$$\begin{aligned}
\operatorname{var}\left(r(s)|r(t)\right) &= \frac{\theta\sigma^2}{2\kappa} + \frac{\sigma^2}{\kappa}\Big(\frac{\theta}{2} - r(t)\Big)e^{-2\kappa(s-t)} \\
&\quad + \frac{\sigma^2}{\kappa}(r(t) - \theta)e^{-\kappa(s-t)} \\
&= \frac{\theta\sigma^2}{2\kappa}\big(1 - e^{-\kappa(s-t)}\big)^2 + r(t)\frac{\sigma^2}{\kappa}\, e^{-\kappa(s-t)}\big(1 - e^{-\kappa(s-t)}\big) \\
&= \frac{\theta(1 - e^{-\kappa(s-t)})}{c} + \frac{2r(t)\, e^{-\kappa(s-t)}}{c} \\
&= \frac{2n}{(2c)^2} + \frac{4\varphi^*}{2c}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \operatorname{var}\left(2c\, r(s)|r(t)\right) &= (2c)^2 \ \operatorname{var}\left(r(s)|r(t)\right) \\
&= 2n + 4(2c\varphi^*)
\end{aligned}$$

So from the definition of the mean and variance of a χ^2 distribution in (2.24) we may conclude that $2c\, r(s)$ has a χ^2 distribution with n degrees of freedom and parameter of non-centrality

$$\varphi = 2c\varphi^* = \frac{4\,\kappa\, r(t)\, e^{-\kappa(s-t)}}{\sigma^2(1 - e^{-\kappa(s-t)})} = 2u$$

To determine the probability density function of $r(s)$, conditional on its value at time t, use the characterisation of the probability density function of a χ^2 distribution in (2.24) with $Y = 2c\,r(s)$:

$$\begin{aligned}
f[r(s)|r(t)] &= 2c\,\frac{e^{-cr(s)}e^{-cr(t)e^{-\kappa(s-t)}}(2cr(s))^q}{2^{q+1}\Gamma(q+1)} \\
&\quad\times\left(1+\frac{1}{2q+2}\frac{4c^2r(s)r(t)e^{-\kappa(s-t)}}{2}\right. \\
&\quad\left.+\frac{1}{(2q+2)(2q+4)}\frac{1}{2!}\left(\frac{4c^2r(s)r(t)e^{-\kappa(s-t)}}{2}\right)^2+\cdots\right) \\
&= \frac{ce^{-cr(s)-cr(t)e^{-\kappa(s-t)}}(cr(s))^q}{\Gamma(q+1)} \\
&\quad\times\left(1+\frac{c^2r(s)r(t)e^{-\kappa(s-t)}}{q+1}+\frac{c^4r^2(s)r^2(t)e^{-2\kappa(s-t)}}{2!(q+1)(q+2)}+\cdots\right) \qquad (2.25)
\end{aligned}$$

This representation of the conditional probability density function can be shown to be equivalent to the formulation used by CIR in their original paper, which makes use of the modified Bessel function[15]:

$$f[r(s)|r(t)] = c\,e^{-u-\nu}\left(\frac{\nu}{u}\right)^{q/2}I_q\left(2\sqrt{u\nu}\right) \qquad (2.26)$$

---

[15] $I_q(\cdot)$ is the modified Bessel function of the first kind of order q. This function is defined as:

$$I_q(z) = \left(\frac{z}{2}\right)^q\sum_{k=0}^{\infty}\frac{\left(\frac{z^2}{4}\right)^k}{k!\Gamma(q+k+1)}$$

and so in the CIR probability density function, we have:

$$\begin{aligned}
I_q\left(2\sqrt{u\nu}\right) &= I_q\left(2c\sqrt{r(s)r(t)}e^{-\frac{1}{2}\kappa(s-t)}\right) \\
&= \left(c\sqrt{r(s)r(t)}e^{-\frac{1}{2}\kappa(s-t)}\right)^q\sum_{k=0}^{\infty}\frac{\left(c^2r(s)r(t)e^{-\kappa(s-t)}\right)^k}{k!\;\Gamma(q+k+1)} \\
&= \left(c^2r(s)r(t)e^{-\kappa(s-t)}\right)^{q/2} \\
&\quad\times\left(\frac{1}{\Gamma(q+1)}+\frac{c^2r(s)r(t)e^{-\kappa(s-t)}}{\Gamma(q+2)}+\frac{c^4r^2(s)r^2(t)e^{-2\kappa(s-t)}}{2!\;\Gamma(q+3)}+\cdots\right)
\end{aligned}$$

Hence, making use of the identity $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, (2.26) expands to:

$$\begin{aligned}
f[r(s)|r(t)] &= ce^{-cr(s)-cr(t)e^{-\kappa(s-t)}}\left(\frac{cr(s)}{cr(t)e^{-\kappa(s-t)}}\right)^{q/2}\left(c^2r(s)r(t)e^{-\kappa(s-t)}\right)^{q/2} \\
&\quad\times\left(\frac{1}{\Gamma(q+1)}+\frac{c^2r(s)r(t)e^{-\kappa(s-t)}}{\Gamma(q+2)}+\frac{c^4r^2(s)r^2(t)e^{-2\kappa(s-t)}}{2!\;\Gamma(q+3)}+\cdots\right) \\
&= \frac{ce^{-cr(s)-cr(t)e^{-\kappa(s-t)}}(cr(s))^q}{\Gamma(q+1)} \\
&\quad\times\left(1+\frac{c^2r(s)r(t)e^{-\kappa(s-t)}}{(q+1)}+\frac{c^4r^2(s)r^2(t)e^{-2\kappa(s-t)}}{2!\;(q+1)(q+2)}+\cdots\right)
\end{aligned}$$

which is equivalent to the distribution in (2.25).

2.7.5. Mean reversion. The conditional mean of the short-term interest rate depicted in (2.19) clearly displays its reverting property.

Consider:

- if $r(t) = \theta$ then $r(s) = \theta$, for all $s \geq t$,
- if $r(t) \neq \theta$ then $\lim_{s\to\infty} r(s) = \theta$,
- for $\kappa \to \infty$ (where κ is the reversion speed), $\mathbb{E}\left[r(s)|r(t)\right] \to \theta$, the long-term mean and $\operatorname{var}\left(r(s)|r(t)\right) \to 0$,
- for $\kappa \to 0$, $\mathbb{E}\left[r(s)|r(t)\right] \to r(t)$, the current rate of interest and $\operatorname{var}\left(r(s)|r(t)\right) \to r(t)\sigma^2(s-t)$.[16]

In the above cases the properties of the future rate of interest are consistent with our expectations, given the structure of the interest rate dynamics.

If $\kappa,\ \theta > 0$, i.e. the interest rate is mean reverting, then as $s \to \infty$, the interest rate approaches the gamma distribution with density function:

$$f[r(\infty)|r(t)] = \frac{\omega^\nu}{\Gamma(\nu)}\, r^{\nu-1}\, e^{-\omega r}$$

where $\omega = 2\kappa/\sigma^2$ and $\nu = 2\kappa\theta/\sigma^2$. The mean and variance are θ and $\sigma^2\theta/2\kappa$ respectively which is also consistent with our expectations of the interest rate dynamics[17].

2.7.6. Negative interest rates. As already identified in §2.7.3 and §2.7.4 $r(t)$ has a χ^2 distribution and so it may be written as a sum of squares of normally distributed random variables:

$$r(t) = \sum_{i=1}^{n} X_i^2(t)$$

Consider the case: $n = 1$, so $r(t) = X_1^2(t)$. Since $X_1(t)$ is normally distributed, we know that for each t, $\mathbb{P}\{r(t) \geq 0\} = 1$. Hence it is also the case that for any $t > 0$, $\mathbb{P}\{\text{There are infinitely many values of } t \text{ where } r(t) = 0\} = 1$.

Now consider $n \geq 2$. For $r(t) = 0,\ t > 0$ we require $X_i(t) = 0$, for every $i = 1, \ldots, n$ and so $\mathbb{P}\{\text{There exists a } t > 0 \text{ such that } r(t) = 0\} = 0$.

---

[16] Apply L'Hopital's Rule to the first term of (2.20) to obtain

$$\lim_{\kappa\to 0} \operatorname{var}\left(r(s)|r(t)\right) = \lim_{\kappa\to 0} r(t)\sigma^2\left(-(s-t)e^{-\kappa(s-t)} + 2(s-t)e^{-2\kappa(s-t)}\right) = r(t)\sigma^2(s-t)$$

[17] This mean and variance can also be obtained by letting $s \to \infty$ in (2.19) and (2.20).

Therefore, to prevent zero (and hence negative) interest rates, we require $n \geq 2$. From the stochastic process describing $r(t)$ in (2.23) we have:

$$\theta = \frac{n\sigma^2}{4\kappa}$$
$$\Rightarrow n = \frac{4\kappa\theta}{\sigma^2}$$

The requirement $n \geq 2$, needed to prevent zero (and hence negative) interest rates translates to:

$$\frac{4\kappa\theta}{\sigma^2} \geq 2$$
$$\Rightarrow 2\kappa\theta \geq \sigma^2$$

2.8. Bond pricing formula

Allowing the bond price to take the form:

$$P(r,t,T) = A(t,T)e^{-B(t,T)r} \tag{2.27}$$

we have[18]:

$$\begin{aligned} P_r &= -ABe^{-Br} \\ P_{rr} &= AB^2e^{-Br} \\ P_t &= A_te^{-Br} - AB_tre^{-Br} \end{aligned}$$

and so the bond price equation (2.18) reduces to:

$$r\left(\tfrac{1}{2}\sigma^2AB^2 + \kappa AB - AB_t + \lambda AB - A\right) = \kappa\theta AB - A_t$$

Since the left-hand side is a function of the short-term interest rate $r(t)$, while the right-hand side is independent of the short-term interest rate, the following two equations must be satisfied:

$$A_t - \kappa\theta AB = 0 \tag{2.28}$$

$$B_t - (\kappa + \lambda)B - \tfrac{1}{2}\sigma^2B^2 + 1 = 0 \tag{2.29}$$

To solve for $A(t,T)$ and $B(t,T)$ first consider (2.29). This is a Ricatti equation[19] with solution $B(t,T) = v(t,T)/u(t,T)$ where $v(t,T)$ and $u(t,T)$ are

[18] Where the subscript indicates a partial derivative.

[19] The general form of the Ricatti equation is:

$$w'(t) + [a(t) + d(t)]\,w(t) + b(t)w^2(t) - c(t) \;=\; 0$$

The solution of this equation can be written as $w(t) = v(t)/u(t)$ where $v(t)$ and $u(t)$ are solutions of the associated system of first order linear equations:

$$\begin{aligned} -v'(t) + c(t)u(t) - d(t)v(t) &= 0 \\ u'(t) - a(t)u(t) - b(t)v(t) &= 0 \end{aligned}$$

For more details on solving the Ricatti equation see [**47**].

solutions to the following system of equations[20]:

$$\begin{aligned} v'(t,T) + u(t,T) - \kappa v(t,T) &= 0 \\ u'(t,T) + \lambda u(t,T) + \tfrac{1}{2}\sigma^2 v(t,T) &= 0 \end{aligned}$$

Set $\tau = T - t$ where T is the bond maturity date; then $\frac{\partial}{\partial t} = -\frac{d}{d\tau}$ and the above system of equations may be written as:

$$-v'(\tau) + u(\tau) - \kappa v(\tau) = 0 \tag{2.30}$$

$$-u'(\tau) + \lambda u(\tau) + \tfrac{1}{2}\sigma^2 v(\tau) = 0 \tag{2.31}$$

From (2.30) we have:

$$u(\tau) = v'(\tau) + \kappa v(\tau) \tag{2.32}$$

$$u'(\tau) = v''(\tau) + \kappa v'(\tau) \tag{2.33}$$

substituting into (2.31) gives:

$$-v''(\tau) - \kappa v'(\tau) + \lambda v'(\tau) + \lambda\kappa v(\tau) + \tfrac{1}{2}\sigma^2 v(\tau) = 0$$

Expressing this in terms of D-operators results in a simple quadratic equation:

$$\left[D^2 - (\lambda - \kappa)D - (\lambda\kappa + \tfrac{1}{2}\sigma^2)\right]v(\tau) = 0$$

The roots of this quadratic equation are $(\gamma + \lambda - \kappa)/2$ and $(-\gamma + \lambda - \kappa)/2$ where $\gamma = \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}$ and hence the solution may be written as:

$$v(\tau) = k_1 e^{(\gamma+\lambda-\kappa)\tau/2} + k_2 e^{(-\gamma+\lambda-\kappa)\tau/2}$$

where k_1 and k_2 are constants. Since $B(T,T) = 0 = v(0)/u(0)$, $v(0) = 0$ and hence $k_1 = -k_2$. Setting $k_1 = 1$ and $k_2 = -1$, $v(\tau)$ becomes:

$$v(\tau) = e^{(\gamma+\lambda-\kappa)\tau/2} - e^{(-\gamma+\lambda-\kappa)\tau/2} \tag{2.34}$$

Substituting (2.34) and

$$v'(\tau) = \tfrac{1}{2}(\gamma + \lambda - \kappa)e^{(\gamma+\lambda-\kappa)\tau/2} - \tfrac{1}{2}(-\gamma + \lambda - \kappa)e^{(-\gamma+\lambda-\kappa)\tau/2}$$

into (2.32) gives:

$$u(\tau) = \tfrac{1}{2}(\gamma + \lambda + \kappa)e^{(\gamma+\lambda-\kappa)\tau/2} - \tfrac{1}{2}(-\gamma + \lambda + \kappa)e^{(-\gamma+\lambda-\kappa)\tau/2} \tag{2.35}$$

Since $\tau = T - t$, the solution of the Riccati equation is obtained from (2.34) and (2.35) as:

[20] v and u are functions of t and T, but T is fixed, hence $v'(t,T)$ and $u'(t,T)$ denote the derivative with respect to t.

$$\begin{aligned}
B(t,T) &= v(\tau)/u(\tau) \\
&= \frac{2\left(e^{(\gamma+\lambda-\kappa)(T-t)/2} - e^{(-\gamma+\lambda-\kappa)(T-t)/2}\right)}{(\gamma+\lambda+\kappa)e^{(\gamma+\lambda-\kappa)(T-t)/2} - (-\gamma+\lambda+\kappa)e^{(-\gamma+\lambda-\kappa)(T-t)/2}} \\
&= \frac{2\left(e^{\gamma(T-t)} - 1\right)}{(\gamma+\lambda+\kappa)e^{\gamma(T-t)} - (-\gamma+\lambda+\kappa)} \\
&= \frac{2\left(e^{\gamma(T-t)} - 1\right)}{(\gamma+\lambda+\kappa)(e^{\gamma(T-t)} - 1) + 2\gamma} \qquad (2.36)
\end{aligned}$$

Now consider equation (2.28) with fixed bond maturity T, so the bond price is a function of t only. Hence:

$$\begin{aligned}
\frac{\partial A}{\partial t} &= \kappa\theta AB \\
\frac{dA}{A} &= \kappa\theta B dt \\
\ln A(t,T) &= -\int_t^T \kappa\theta B(s,T)ds \\
A(t,T) &= \exp\left(-\kappa\theta\int_t^T B(s,T)ds\right) \qquad (2.37)
\end{aligned}$$

where κ and θ are constants. Substituting $B(t,T)$ from (2.36) into (2.37):

$$\begin{aligned}
A(t,T) &= \exp\left(-\kappa\theta\int_t^T B(s,T)\,ds\right) \\
&= \exp\left(-2\kappa\theta\int_t^T \frac{e^{\gamma(T-s)} - 1}{(\gamma+\lambda+\kappa)(e^{\gamma(T-s)} - 1) + 2\gamma}\,ds\right)
\end{aligned}$$

Let $y = e^{\gamma(T-s)}$ then $\frac{dy}{ds} = -\gamma e^{\gamma(T-s)}$ and $ds = -\frac{dy}{\gamma e^{\gamma(t-s)}} = -\frac{dy}{\gamma y}$. Making this substitution and noting that $(\gamma-\lambda-\kappa)(\gamma+\lambda+\kappa) = \gamma^2 - (\kappa+\lambda)^2 = 2\sigma^2$, the integral in the above equation becomes:

$$\int_t^T \frac{e^{\gamma(T-s)} - 1}{(\gamma+\lambda+\kappa)(e^{\gamma(T-s)} - 1) + 2\gamma}\,ds$$

$$= \frac{1}{\gamma}\int_{e^{\gamma(T-t)}}^1 \frac{-(y-1)}{(\gamma+\lambda+\kappa)(y-1) + 2\gamma}\,\frac{dy}{y}$$

$$= \frac{1}{\gamma}\int_{e^{\gamma(T-t)}}^1 \left[\frac{-2\gamma/(\gamma-\lambda-\kappa)}{(\gamma+\lambda+\kappa)(y-1) + 2\gamma} + \frac{1}{(\gamma-\lambda-\kappa)}\frac{1}{y}\right] dy$$

$$
\begin{aligned}
&= \frac{-2}{(\gamma-\lambda-\kappa)(\gamma+\lambda+\kappa)} \ln\left[(\gamma+\lambda+\kappa)(y-1)+2\gamma\right]\Big|_{y=e^{\gamma(T-t)}}^{y=1} \\
&\qquad + \frac{1}{\gamma(\gamma-\lambda-\kappa)} \ln y \Big|_{y=e^{\gamma(T-t)}}^{y=1} \\
&= \frac{1}{\sigma^2}\left[-\ln\big((\gamma+\lambda+\kappa)(y-1)+2\gamma\big) + \frac{(\gamma+\lambda+\kappa)}{2\gamma}\ln y\right]\Bigg|_{y=e^{\gamma(T-t)}}^{y=1} \\
&= \frac{1}{\sigma^2}\left[\ln \frac{y^{(\gamma+\lambda+\kappa)/2\gamma}}{(\gamma+\kappa+\lambda)(y-1)+2\gamma}\right]\Bigg|_{y=e^{\gamma(T-t)}}^{y=1} \\
&= \frac{1}{\sigma^2}\left[-\ln 2\gamma - \ln \frac{e^{(\gamma+\lambda+\kappa)(T-t)/2}}{(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma}\right]
\end{aligned}
$$

and hence the solution for $A(t,T)$ is:

$$
\begin{aligned}
A(t,T) &= \exp\left(\frac{2\kappa\theta}{\sigma^2}\ln\frac{2\gamma e^{(\gamma+\lambda+\kappa)(T-t)/2}}{(\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma}\right) \\
&= \left(\frac{2\gamma e^{(\gamma+\lambda+\kappa)(T-t)/2}}{(\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma}\right)^{2\kappa\theta/\sigma^2} \qquad (2.38)
\end{aligned}
$$

An analysis similar to that in the derivation of Vasicek's model in Chapter 1 can be applied to determine the bond price dynamics. Since the bond price is a function of the short-term interest rate, Ito's Lemma is used to give:

$$dP = \mu_B P dt - \sigma_B P dz$$

where

$$
\begin{aligned}
\sigma_B &= -\frac{1}{P}\left(\sigma\sqrt{r}\frac{\partial P}{\partial r}\right) \\
&= -\frac{e^{Br}}{A}\,\sigma\sqrt{r}\,ABe^{-Br} \\
&= \sigma\sqrt{r}B
\end{aligned}
$$

The existence of a factor risk premium $q(r,t)$, implies:

$$\mu_B - r = q(r,t)\,\sigma_B$$

Vasicek assumes a constant market price of risk, while CIR specify:

$$
\begin{aligned}
q(r,t) &= -\frac{\lambda\sqrt{r}}{\sigma}, \qquad \lambda = \text{constant} \\
\Rightarrow \mu_B &= r - \lambda r B
\end{aligned}
$$

and hence under the CIR model the bond price process is specified by:

$$dP = r\left(1 - \lambda B\right) Pdt - B\sigma\sqrt{r}dz$$

2.9. Properties of the bond price under the CIR model

2.9.1. Yield-to-maturity. Since it is market convention to quote bond prices in terms of yield-to-maturity, it is more insightful to examine the behaviour of the yield-to-maturity in the case of a very short and very long time to maturity. Since a zero coupon bond is a pure discount instrument, its price is written as:

$$P(r,t,T) = e^{-(T-t)R(r,t,T)} \tag{2.39}$$

where $R(r,t,T)$ is the yield-to-maturity. Equating (2.27) and (2.39) we derive the yield-to-maturity in terms of $A(t,T)$ and $B(t,T)$ as:

$$\begin{aligned} A(t,T)e^{-B(t,T)r} &= e^{-(T-t)R(r,t,T)} \\ \ln A(t,T) - B(t,T)r &= -(T-t)R(r,t,T) \\ R(r,t,T) &= \frac{rB(t,T) - \ln A(t,T)}{T-t} \end{aligned}$$

As $t \to T$, $R(r,t,T) \to r$,[21]since as the bond approaches maturity, it converges to an instrument with instantaneous maturity. Now consider the yield-to-maturity as $T \to \infty$. This may be viewed as the yield on a perpetual bond:

$$\begin{aligned} R(r,t,\infty) &= \lim_{T\to\infty} R(r,t,T) = \lim_{T\to\infty} \frac{-\ln P(r,t,T)}{T-t} \\ &= \lim_{T\to\infty} \frac{-\ln A(t,T) + rB(t,T)}{T-t} \end{aligned}$$

Consider $\ln A(t,T)$ where $A(t,T)$ is given by (2.38):

$$\begin{aligned} \ln A(t,T) &= \frac{2\kappa\theta}{\sigma^2}\Big[\ln 2\gamma e^{(\gamma+\lambda+\kappa)(T-t)/2} \\ &\quad - \ln\left[(\gamma+\lambda+\kappa)(e^{\gamma(T-t)} - 1) + 2\gamma\right]\Big] \\ &= \frac{2\kappa\theta}{\sigma^2}\Big[(\gamma+\lambda+\kappa)(T-t)/2 \\ &\quad + \ln 2\gamma - \ln\left[(\gamma+\lambda+\kappa)(e^{\gamma(T-t)} - 1) + 2\gamma\right]\Big] \end{aligned}$$

[21] e^x can be approximated by its power series expansion as:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Now, since

$$\begin{aligned}
&\lim_{T\to\infty}\left[e^{\gamma(T-t)}-1\right]=\lim_{T\to\infty}\left[e^{\gamma(T-t)}\right]\\
&\lim_{T\to\infty}\left[(\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma\right]=\lim_{T\to\infty}\left[(\gamma+\lambda+\kappa)\,e^{\gamma(T-t)}\right]\\
&\lim_{T\to\infty}\left[(\gamma+\lambda+\kappa)(T-t)/2+\ln 2\gamma\right]=\lim_{T\to\infty}\left[(\gamma+\lambda+\kappa)(T-t)/2\right]
\end{aligned}$$

we have:

$$\begin{aligned}
\lim_{T\to\infty}\ln A(t,T) &= \lim_{T\to\infty}\frac{2\kappa\theta}{\sigma^2}\left[(\gamma+\lambda+\kappa)(T-t)/2-\ln(\gamma+\lambda+\kappa)\,e^{\gamma(T-t)}\right]\\
&= \lim_{T\to\infty}\frac{\kappa\theta}{\sigma^2}\left[(\gamma+\lambda+\kappa)(T-t)-2\gamma(T-t)\right]\\
&= \lim_{T\to\infty}\frac{\kappa\theta}{\sigma^2}(-\gamma+\lambda+\kappa)(T-t)
\end{aligned}$$

Also from (2.36):

$$\lim_{T\to\infty}B(t,T)=\lim_{T\to\infty}\frac{2(e^{\gamma(T-t)}-1)}{(\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma}=\frac{2}{\gamma+\lambda+\kappa}$$

---

and hence [**45**]

$$\begin{aligned}
e^{(\gamma+\lambda+\kappa)(T-t)/2} &\approx 1+\frac{(\gamma+\lambda+\kappa)(T-t)}{2}+\frac{(\gamma+\lambda+\kappa)^2(T-t)^2}{8}\\
e^{\gamma(T-t)} &\approx 1+\gamma(T-t)+\frac{\gamma^2(T-t)^2}{2}
\end{aligned}$$

Therefore as $t\to T$ we have:

$$\begin{aligned}
e^{(\gamma+\lambda+\kappa)(T-t)/2} &\to 1+\frac{(\gamma+\lambda+\kappa)(T-t)}{2}\\
\Rightarrow A(t,T) &\to \left[\frac{2\gamma\,(1+(\gamma+\lambda+\kappa)(T-t)/2)}{\gamma(\gamma+\lambda+\kappa)(T-t)+2\gamma}\right]^{2\kappa\theta/\sigma^2}\\
&= \left[\frac{2\gamma+\gamma(\gamma+\lambda+\kappa)(T-t)}{2\gamma+\gamma(\gamma+\lambda+\kappa)(T-t)}\right]^{2\kappa\theta/\sigma^2}\\
&= 1
\end{aligned}$$

and also

$$\begin{aligned}
e^{\gamma(T-t)} &\to 1+\gamma(T-t)\\
\Rightarrow B(t,T) &\to \frac{2\gamma(T-t)}{\gamma(\gamma+\lambda+\kappa)(T-t)+2\gamma}\\
&= \frac{T-t}{1+(\gamma+\lambda+\kappa)(T-t)/2}\\
&\approx T-t
\end{aligned}$$

for small $T-t$. Hence from (2.27), $P(r,t,T)=e^{-r(T-t)}$.

and hence the yield on a perpetual bond is[22]:

$$
\begin{aligned}
R(r,t,\infty) &= \lim_{T\to\infty} \frac{B(t,T)r - \ln A(t,T)}{T-t} \\
&= \lim_{T\to\infty} \frac{2r/(\gamma+\lambda+\kappa) - \frac{\kappa\theta}{\sigma^2}(-\gamma+\lambda+\kappa)(T-t)}{T-t} \\
&= \lim_{T\to\infty} \frac{2r + 2\kappa\,\theta(T-t)}{(\gamma+\lambda+\kappa)(T-t)} \\
&= \lim_{T\to\infty} \Big[\frac{2r}{(\gamma+\lambda+\kappa)(T-t)} + \frac{2\kappa\,\theta}{\gamma+\lambda+\kappa}\Big] \\
&= \frac{2\kappa\,\theta}{\gamma+\lambda+\kappa}
\end{aligned}
$$

Hence, for bonds with increasing maturity, the yield approaches a limit independent of current rate of interest, but proportional to the mean reversion level [**45**].

2.9.2. Possible shapes of the term structure. As for the Vasicek model, the CIR term structure can assume various shapes according to the level of the current interest rate, $r(t)$. See Figure 2.1 below. For $r(t) < \frac{2\kappa\theta}{\gamma+\lambda+\kappa}$, the long-term yield, the term structure is uniformly increasing while for $r(t) > \frac{2\kappa\theta}{\kappa+\lambda}$ the term structure is uniformly decreasing. For values of $r(t)$ lying between these two extremes the term structure is humped.

2.10. Extending the model to allow time-dependent drift

Consider extending the model specification to allow for a time-dependent drift parameter. Hence the short-term interest rate dynamics become:

$$dr = \kappa(\theta(t) - r)dt + \sigma\sqrt{r}dz(t) \tag{2.40}$$

Due to the Markovian nature of the model, we assume that all information about past movements and expectations of future movements is contained in the current (observed) term structure. Therefore the functional form of the time-dependent parameter $\theta(t)$, may be determined from observed bond prices and the values of the constant parameters. No prior restrictions are placed on the functional form of $\theta(t)$ since it is determined so as to reflect the specific observed term structure.

Consider the conditional expectation of $r(s)$ with the time-dependent parameter $\theta(t)$. Following the methodology of §(2.7.1) we have the integral form

[22]Here make use of the following:

$$(\gamma+\lambda+\kappa)(-\gamma+\lambda+\kappa) = (\kappa+\lambda)^2 - \gamma^2 = -2\sigma^2$$

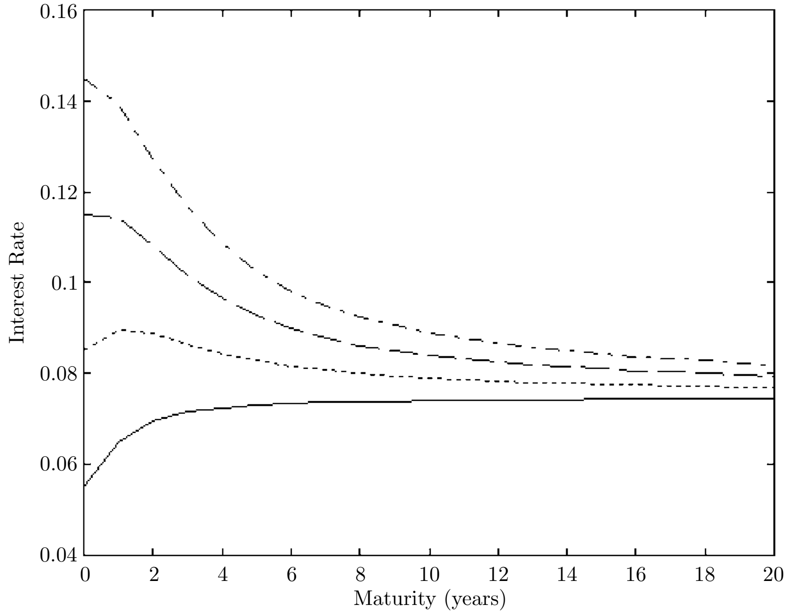

FIGURE 2.1. Possible shapes of the term structure. $\kappa = 0.3$, $\lambda = 0$, $\sigma = 0.6$, $\theta = 0.15$

of the short-term interest rate process:

$$r(s) = r(t) + \kappa \int_t^s \big(\theta(u) - r(u)\big)\, du + \sigma \int_t^s \sqrt{r(u)}\, dz(u)$$

Taking expectations and differentiating with respect to s produces:

$$\begin{aligned}
\mathbb{E}\left[r(s)|r(t)\right] &= r(t) + \kappa \int_t^s \big(\theta(u) - \mathbb{E}\left[r(u)|r(t)\right]\big)\, du \\
\frac{\partial\, \mathbb{E}\left[r(s)|r(t)\right]}{\partial s} &= \kappa\big(\theta(s) - \mathbb{E}\left[r(s)|r(t)\right]\big) \\
\Rightarrow \frac{\partial}{\partial s}\big(e^{\kappa(s-t)}\, \mathbb{E}\left[r(s)|r(t)\right]\big) &= \kappa\, e^{\kappa(s-t)}\, \mathbb{E}\left[r(s)|r(t)\right] + e^{\kappa(s-t)} \frac{\partial\, \mathbb{E}\left[r(s)|r(t)\right]}{\partial s} \\
&= \kappa\, \theta(s)\, e^{\kappa(s-t)}
\end{aligned}$$

so, integrating over $[t, s]$ gives the conditional expectation of $r(s)$ as:

$$e^{\kappa(s-t)}\, \mathbb{E}\left[r(s)|r(t)\right] - r(t) = \kappa \int_t^s \theta(u)\, e^{\kappa(u-t)}\, du$$

$$\Rightarrow\ \mathbb{E}\left[r(s)|r(t)\right] = r(t)\, e^{-\kappa(s-t)} + \kappa \int_t^s \theta(u)\, e^{-\kappa(s-u)}\, du \tag{2.41}$$

The bond price takes the same functional form as specified in (2.27), with a modification to one of the parameters as depicted below:

$$P(r,t,T) = \widehat{A}(t,T)e^{-B(t,T)r} \tag{2.42}$$

where:

$$B(t,T) = \frac{2\left(e^{\gamma(T-t)}-1\right)}{(\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma} \tag{2.43}$$

$$\widehat{A}(t,T) = \exp\left(-\kappa\int_t^T \theta(s)B(s,T)\,ds\right) \tag{2.44}$$

Given this formulation of the bond price and the observed term structure, (2.44) can be solved for $\theta(s)$ for all $s \in [t,T]$ which could then be used in conjunction with (2.41) to determine future expectations of the short-term interest rate as specified by the current observed term structure.

2.11. Comparison of the Vasicek and CIR methods of derivation

The Vasicek and CIR models are very similar in their structure[23] and tractability, but their key difference lies in the derivation. Vasicek enforces the no arbitrage requirement between bonds but does not consider the existence of an underlying equilibrium economy consistent with the model. CIR begin with a specification of the equilibrium economy, from within which they obtain the valuation model. The following factors are contained within the CIR economy:

- variables affecting the bond price,
- endogenously determined stochastic properties driving the underlying variables,
- the form of the factor risk premia.

Vasicek makes assumptions about the variables affecting the bond price and the stochastic factors driving these variables. These assumptions are exogenously specified and imposed directly on the relevant variables. Consider these assumptions in the framework of the CIR model:

(1) the bond price is assumed to be determined by the short-term interest rate only,
(2) the short-term interest rate r, is assumed to follow the stochastic process
$dr = \kappa(\theta^* - r)dt + \sigma\sqrt{r}dz$

Application of Ito's Lemma and existence of the risk premium determines the excess expected return on a bond, that is $\mu(t,T) - r =$ excess expected return $= \Upsilon(r,t,T)$

$$\Rightarrow \tfrac{1}{2}\sigma^2 rP_{rr} + \kappa(\theta^* - r)P_r + P_t - rP = \Upsilon(r,t,T) \tag{2.45}$$

[23]They apply slightly different functional forms to the volatility of the short-term interest rate and the market price of risk.

If there exists an underlying equilibrium economy which supports (1) and (2), then this function $\Upsilon(r,t,T)$ must exist. However, its dependence on the underlying variables is unspecified.

To preclude arbitrage Υ must take on the following form:

$$\Upsilon(r,t,T) = \Psi(r,t)P_r(r,t,T) \tag{2.46}$$

where $\Psi(r,t)$ is the required risk premium. Not all functions $\Upsilon(r,t,T)$ will satisfy (2.45) and (2.46) and hence definite restrictions are placed on the functional form of the excess return.

However, this approach to the specification of a complete model of the term structure may lead to problems:

(1) Assumptions (1) and (2) do not guarantee a consistent underlying equilibrium economy;
(2) The no arbitrage approach does not guarantee the absence of arbitrage for every choice of $\Psi(r,t)$.

The model specified by CIR does have a consistent underlying equilibrium economy and hence precludes arbitrage. Consider the following example which does not meet all the requirements specified by the CIR model and hence leads to disequilibrium in the underlying economy. Assuming $\Psi(r,t) = \Psi_0 + \lambda r$, (2.45) becomes

$$\begin{aligned} \tfrac{1}{2}\,\sigma^2 r P_{rr} + \kappa(\theta^* - r)P_r + P_t - rP - (\Psi_0 + \lambda r)\,P_r &= 0 \\ \Rightarrow \tfrac{1}{2}\,\sigma^2 r P_{rr} + \kappa\Big(\theta^* - \frac{\Psi_0}{\kappa} - r\Big)P_r + P_t - rP - \lambda r P_r &= 0 \end{aligned} \tag{2.47}$$

This is the same as (2.18) with $\theta = \theta^* - \frac{\Psi_0}{\kappa}$, so the bond price takes the form:

$$P(r,t,T) = A(t,T)' e^{-B(t,T)r}$$

where

$$\begin{aligned} A(t,T)' &= \left(\frac{2\gamma e^{(\kappa+\gamma+\lambda)(T-t)/2}}{(\kappa+\gamma+\lambda)\left(e^{\gamma(T-t)}-1\right)+2\gamma}\right)^{2\kappa\left(\theta^* - \frac{\Psi_0}{\kappa}\right)/\sigma^2} \\ &= \left(\frac{2\gamma e^{(\kappa+\gamma+\lambda)(T-t)/2}}{(\kappa+\gamma+\lambda)\left(e^{\gamma(T-t)}-1\right)+2\gamma}\right)^{\frac{2\kappa\theta^*}{\sigma^2}\,\frac{(\kappa\theta^* - \Psi_0)}{\kappa\theta}} \\ &= A(t,T)^{(\kappa\theta^* - \Psi_0)/\kappa\theta} \end{aligned}$$

The solution of the bond price equation (2.47) becomes:

$$P(r,t,T) = A(t,T)^{(\kappa\theta^* - \Psi_0)/\kappa\theta}\, e^{-rB(t,T)}$$

and the bond price process may be specified as:

$$\begin{aligned} dP(r,t,T) &= (r - (\Psi_0 + \lambda r)\,B(t,T))\,P(r,t,T)\,dt \\ &\quad -B(t,T)\,\sigma\sqrt{r}\,P(r,t,T)\,dz \end{aligned} \tag{2.48}$$

The linear form of the risk premium chosen above satisfies the no arbitrage condition and appears advantageous for empirical studies, but it can easily be shown that the resulting model is in fact not viable. Consider $r = 0$. Since the bond is instantaneously riskless, it should over the next instant, yield the corresponding risk-free rate. However, the bond price dynamics (2.48) reduce to:

$$dP(r,t,T) = -\Psi_0 B(t,T) P(r,t,T)\,dt \tag{2.49}$$

and hence the instantaneous rate of return differs from the prevailing risk-free rate and the model guarantees arbitrage opportunities instead of precluding them. This model breaks down because there is no underlying economic equilibrium which is consistent with the chosen risk premium.

2.12. More complicated model specifications

The specific term structure model derived by CIR assumes the state of technology is represented by a single state variable, and randomness within the economy is explained by the stochastic dynamics of this variable. Hence bond prices of all maturities are determined by a single random variable, the short-term interest rate. The model does allow some flexibility, since the term structure may assume a number of shapes, but the nature of single factor models implies that price changes in bonds of all maturities are perfectly correlated and independent of the path followed by the short-term interest rate to reach its current value.

Multi-factor models which allow a richer specification of the technology introduce more flexibility into the term structure, but often this is accompanied by an undesirable increase in complexity and lack of analytical tractability. The two models considered thus far always involve an explanatory variable that is not directly observable in the market. This is the market price of risk or factor risk premium. It is dependent on the utility function of individual investors which cannot be empirically determined. Multi-factor models will tend to have even more investor-specific, and hence unobservable, variables. At times, it may be possible to express these unobservable variables as functions of the endogenously determined prices (e.g. the risk-free rate of interest) and thereby eliminate them from the pricing model. This is the case with the Brennan and Schwartz (1979) model discussed in the next chapter.

2.13. Conclusion

The main characteristic of the CIR model is that current prices and stochastic properties of all contingent claims are derived endogenously. Since CIR use the rational asset pricing model to determine the term structure of interest rates, the following factors are all material in the derivation: investor anticipations, risk aversion, available investment alternatives and preferences with respect to timing of consumption. Equilibrium asset pricing principles are combined with appropriate models of stochastic processes describing the evolution of randomness in the economy, to derive consistent and possibly refutable theories.

The drawback of the CIR model is that it is only a general equilibrium model within their simplified and stylised economy. The investor-specific utility function always enters the model via the market price of risk; for calibration purposes it must be empirically estimated. Model risk will arise since reality is being forced into a simplified model. The following view may be taken of the CIR model:

(1) if the specification of the economic model is correct,
(2) if the stochastic process chosen for the short-term interest rate is in fact the 'true' process describing its development,
(3) if the investor's utility function is fully specified,

then we may say that the model is fully specified and the endogenously derived term structure is the observed term structure. However, none of the above specifications is known with any certainty. In particular, the utility function is difficult to determine empirically. Hence, option prices and other factors derived from the CIR model cannot be seen as accurate quantifications of market characteristics, but rather as descriptive qualifications.

Later models assume that the general structure of the yield curve dynamics are known *a priori*. Details of various parameters, which are obtained from more fundamental factors[24], are left unspecified. The current yield curve is used to fit the unspecified parameters. If the assumed structure of the model is an accurate description of reality then the unobservable quantities can be determined in this manner. Certain discrepancies will exist between observed market rates and those derived from the model. An analysis of these discrepancies will reveal their importance and effect as well as indicating whether they are the result of a poor model or a violation of the initial assumptions. Neither the Vasicek nor CIR models allows for complex yield curve patterns and hence tend to be poor representations of observed yield curves.

[24] Such as the utility function.

CHAPTER 3

The Brennan and Schwartz Model

Brennan and Schwartz (BS) [**10**] challenge the primary assumption of many models. That is: all information about future interest rates is contained in the current instantaneous short-term interest rate and hence the prices of all default-free bonds may be represented as time-dependent functions of this instantaneous rate only. They point out that this is not an accurate representation of reality and propose an interest rate model based on the assumption that the whole term structure can be expressed as a function of the yields of the longest and shortest maturity default-free bonds.

They incorporate the assumption that the long-term rate of interest contains information about future values of the short-term interest rate. This long-term interest rate is the second, exogenously specified variable and the term structure between the short- and long-term interest rates is modelled as a function of these two rates. Other models derive the whole term structure, and hence the long-term rate of interest as a function of the short-term interest rate.

3.1. The generic model

Let

$r(t)$	–	instantaneous rate of interest,
$l(t)$	–	long-term interest rate represented by the yield on a consol bond paying a continuous dividend.

BS [**10**] assume $r(t)$ and $l(t)$ to follow a joint Gauss–Markov stochastic process of the general form:

$$dr = \beta_1(r,l,t)\,dt + \eta_1(r,l,t)\,dz_1 \tag{3.1}$$

$$dl = \beta_2(r,l,t)\,dt + \eta_2(r,l,t)\,dz_2 \tag{3.2}$$

where t is the current time and

$dz_1,\ dz_2$	–	represent Wiener processes[1],
$\beta_1(\cdot),\ \beta_2(\cdot)$	–	are the expected, instantaneous rates of change of r and l respectively,

$\eta_1^2(\cdot),\ \eta_2^2(\cdot)$ – are the instantaneous variances of r and l respectively,

ρ – instantaneous correlation between the unanticipated changes in r and l i.e. $dz_1 dz_2 = \rho\, dt$

Let $B(r, l, \tau)$ be the price of a zero coupon bond with maturity date $T = t + \tau$ and unit maturity value. Since this bond price is a function of the two rates of interest and time to maturity τ, we may apply Ito's Lemma to derive the equation of the stochastic process for the bond price[2]:

$$
\begin{aligned}
dB &= \frac{\partial B}{\partial r}\, dr + \frac{\partial B}{\partial l}\, dl + \frac{\partial B}{\partial \tau}\, d\tau + \frac{1}{2}\frac{\partial^2 B}{\partial r^2}\, dr\, dr + \frac{1}{2}\frac{\partial^2 B}{\partial l^2}\, dl\, dl + \frac{\partial^2 B}{\partial r \partial l}\, dr\, dl \\
&= \frac{\partial B}{\partial r}(\beta_1\, dt + \eta_1\, dz_1) + \frac{\partial B}{\partial l}(\beta_2\, dt + \eta_2\, dz_2) - \frac{\partial B}{\partial \tau}\, dt + \frac{1}{2}\frac{\partial^2 B}{\partial r^2}\, \eta_1^2\, dt \\
&\quad + \frac{1}{2}\frac{\partial^2 B}{\partial l^2}\, \eta_2^2\, dt + \frac{\partial^2 B}{\partial r \partial l}\, \eta_1\, \eta_2\, \rho\, dt
\end{aligned}
$$

and hence the bond price process may be written as

$$
\frac{dB}{B} = \mu(r, l, \tau)\, dt + \frac{1}{B}\frac{\partial B}{\partial r}\, \eta_1\, dz_1 + \frac{1}{B}\frac{\partial B}{\partial l}\, \eta_2\, dz_2 \tag{3.3}
$$

where

$$
\begin{aligned}
\mu(r, l, \tau) = \frac{1}{B}\Bigg(&\beta_1 \frac{\partial B}{\partial r} + \beta_2 \frac{\partial B}{\partial l} + \frac{1}{2}\frac{\partial^2 B}{\partial r^2}\eta_1^2 dt \\
&+ \frac{1}{2}\frac{\partial^2 B}{\partial l^2}\eta_2^2 dt + \frac{\partial^2 B}{\partial r \partial l}\eta_1 \eta_2 \rho dt - \frac{\partial B}{\partial \tau}\Bigg)
\end{aligned} \tag{3.4}
$$

An arbitrage argument is applied to derive the equilibrium relationship between bonds of various maturities. This relationship takes the form of a partial differential equation which places constraints on the risk premia of bonds of various maturities. We consider a portfolio of three bonds of different maturities. This allows us to hedge away the uncertainty associated with both stochastic variables, i.e. r and l. Let:

P – portfolio of bonds of three different maturities,

x_i – amount invested in bond with maturity $\tau_i,\ i = 1, 2, 3$

---

[1] As usual, the Wiener processes dz_1 and dz_2 have the following characteristics:

- $\mathbb{E}[dz_1] = \mathbb{E}[dz_2] = 0$,
- $dz_1^2 = dz_2^2 = dt$.

This last point is not technically precise, but is used as an informal representation of the discrete time formulation:

$$
\big(z(t_{k+1}) - z(t_k)\big)^2 \approx t_{k+1} - t_k
$$

where z is some Wiener process and $t_k,\ t_{k+1}$ are consecutive points in time. See [**49**].

[2] Since $\tau = T - t$ we have $\frac{\partial}{\partial \tau} = -\frac{\partial}{\partial t}$

with $\sum_i x_i = 1$. Ito's Lemma is applied to derive the rate of return on the portfolio as:

$$\begin{aligned}\frac{dP}{P} &= [x_1\mu(\tau_1) + x_2\mu(\tau_2) + x_3\mu(\tau_3)]\,dt \\ &+ \left[x_1\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + x_2\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + x_3\frac{1}{B}\frac{\partial B}{\partial r}\eta_1\right]dz_1 \\ &+ \left[x_1\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 + x_2\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 + x_3\frac{1}{B}\frac{\partial B}{\partial l}\eta_2\right]dz_2\end{aligned}$$

This return can be made instantaneously deterministic by setting the coefficients of the Wiener processes equal to zero:

$$x_1\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + x_2\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + x_3\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 = 0 \tag{3.5}$$

$$x_1\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 + x_2\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 + x_3\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 = 0 \tag{3.6}$$

To preclude arbitrage profits, the portfolio return must be the risk-free rate of interest. Hence:

$$x_1\mu(\tau_1) + x_2\mu(\tau_2) + x_3\mu(\tau_3) = r$$

that is:

$$x_1\,[\mu(\tau_1) - r] + x_2\,[\mu(\tau_2) - r] + x_3\,[\mu(\tau_3) - r] = 0 \tag{3.7}$$

Equations (3.5), (3.6), (3.7) form a system of three linear, homogeneous equations in three unknowns. This system is consistent only if:

$$\begin{aligned}\mu(\tau_1) - r &= \lambda_1(r,l,t)\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + \lambda_2(r,l,t)\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 \\ \mu(\tau_2) - r &= \lambda_1(r,l,t)\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + \lambda_2(r,l,t)\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 \\ \mu(\tau_3) - r &= \lambda_1(r,l,t)\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + \lambda_2(r,l,t)\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2\end{aligned}$$

and hence this relationship holds for any bond maturity τ, and may be written as:

$$\mu(\tau) - r = \lambda_1(r,l,t)\frac{1}{B}\frac{\partial B}{\partial r}\,\eta_1 + \lambda_2(r,l,t)\frac{1}{B}\frac{\partial B}{\partial l}\,\eta_2 \tag{3.8}$$

The functions $\lambda_1(r,l,t)$ and $\lambda_2(r,l,t)$ are independent of the bond maturity. The return on a bond in excess of the risk-free rate of interest is the premium required to compensate the investor for the additional risk. Equation (3.8) expresses this instantaneous risk premium on a discount bond as the sum of two factors. These factors are proportional to the partial covariances of the bond return with the unanticipated changes in each of the two exogenous variables. These partial covariances are represented by $\frac{1}{B}\frac{\partial B}{\partial r}\eta_1$ and $\frac{1}{B}\frac{\partial B}{\partial l}\eta_2$ respectively. The proportionality factors $\lambda_1(r,l,t)$ and $\lambda_2(r,l,t)$, may be interpreted as the market prices of instantaneous and long-term interest rate risk. These market prices of risk are investor-specific and depend on the investors' utility functions.

Substituting (3.4) into (3.8) leads to the partial differential equation for the price of a discount bond:

$$\begin{aligned} \frac{\partial B}{\partial r}\,(\beta_1-\lambda_1\eta_1)+\frac{\partial B}{\partial l}\,(\beta_2-\lambda_2\eta_2)+\frac{1}{2}\frac{\partial^2 B}{\partial r^2}\,\eta_1^2+\frac{1}{2}\frac{\partial^2 B}{\partial l^2}\,\eta_2^2 \\ +\frac{\partial^2 B}{\partial r\partial l}\,\rho\,\eta_1\eta_2-\frac{\partial B}{\partial \tau}-rB=0 \end{aligned} \tag{3.9}$$

This bond pricing equation is dependent on two utility specific variables, i.e. the two market prices of risk. Assuming l to be the yield on a consol bond allows us to eliminate the market price of long-term interest rate risk $\lambda_2(r,l,t)$. A consol bond is a bond of infinite maturity paying a continuous coupon of \$1 per annum. Let $V(l)$ be the price of this consol bond, then:

$$V(l)=l^{-1} \tag{3.10}$$

Applying Ito's Lemma to $V(l)$ gives:

$$\begin{aligned} dV &= \frac{\partial V}{\partial l}\,dl+\frac{1}{2}\frac{\partial^2 V}{\partial l^2}\,dl\,dl \\ &= -\frac{1}{l^2}\,(\beta_2\,dt+\eta_2\,dz_2)+\frac{1}{l^3}\,\eta_2^2\,dt \end{aligned}$$

and so:

$$\frac{dV}{V}=\left(\frac{\eta_2^2}{l^2}-\frac{\beta_2}{l}\right)dt-\left(\frac{\eta_2}{l}\right)dz_2 \tag{3.11}$$

where $-\left(\frac{\eta_2}{l}\right)$ is the partial covariance of the bond's instantaneous rate of return with the unanticipated changes in l. The instantaneous rate of return on the consol bond $\mu(\infty)$, consists of the capital gain and the rate of coupon payment; hence:

$$\mu(\infty)=\frac{\eta_2^2}{l^2}-\frac{\beta_2}{l}+l$$

Since we are in the risk-free world the return on the consol must be the risk-free rate of interest and hence the equilibrium risk premium relationship (3.8) must be satisfied:

$$\mu(\infty)-r \;=\; \lambda_2(r,l,t)\frac{1}{V}\frac{\partial V}{\partial l}\,\eta_2$$

$$
\begin{aligned}
\frac{\eta_2^2}{l^2} - \frac{\beta_2}{l} + l - r &= \lambda_2(r,l,t)\left(-\frac{\eta_2}{l}\right) \\
\Rightarrow \lambda_2(r,l,t) &= -\frac{\eta_2}{l} + \frac{\beta_2 - l^2 + r\,l}{\eta_2}
\end{aligned}
\tag{3.12}
$$

Here we have expressed the market price of long-term interest rate risk as a function of r, l and the parameters of the stochastic process of l. This allows us to reduce the number of utility-dependent factors in the bond price equation. Substituting (3.12) into the bond price equation (3.9) results in:

$$
\begin{aligned}
\frac{1}{2}\frac{\partial^2 B}{\partial r^2}\,\eta_1^2 + \frac{1}{2}\frac{\partial^2 B}{\partial l^2}\,\eta_2^2 + \frac{\partial^2 B}{\partial r \partial l}\,\rho\,\eta_1\,\eta_2 + \frac{\partial B}{\partial r}\,(\beta_1 - \lambda_1\eta_1) \\
+\frac{\partial B}{\partial l}\left(l^2 - rl + \frac{\eta_2^2}{l}\right) - \frac{\partial B}{\partial \tau} - rB = 0
\end{aligned}
\tag{3.13}
$$

which is independent of the market price of long-term interest rate risk, λ_2. Applying the boundary condition $B(r,l,0) = 1$, we may solve the above equilibrium bond pricing equation for any maturity. Since the entire term structure of interest rates may be inferred from these bond prices, we conclude that the term structure, at any point in time, is a function of r, l and λ_1.

By choosing l to be the yield on a consol bond, we are able to determine the bond price independent of λ_2 and β_2. Since the consol is a traded security, the risk associated with l may be hedged away. This is the same as the result obtained by Black and Scholes in deriving the option pricing equation, which is independent of the expected return on the underlying asset. BS [**10**] make the observation that the number of investor-specific (utility-dependent) variables in the pricing equation is equal to the number of state variables (excluding time), less the number of variables for which all the partial derivatives are known. Knowledge of the partial derivatives allows the associated risk to be hedged away. This phenomenon may be demonstrated more generally as follows: the partial derivatives of the consol bond (which represents the long-term interest rate), are known. Also, since it is a traded security, it must satisfy the bond pricing equation (3.9) [**51**], hence:

$$
\begin{aligned}
\frac{\partial V}{\partial r}\,(\beta_1 - \lambda_1\eta_1) + \frac{\partial V}{\partial l}\,(\beta_2 - \lambda_2\eta_2) + \frac{1}{2}\frac{\partial^2 V}{\partial r^2}\,\eta_1^2 + \frac{1}{2}\frac{\partial^2 V}{\partial l^2}\,\eta_2^2 \\
+\frac{\partial^2 V}{\partial r \partial l}\,\rho\,\eta_1\,\eta_2 - \frac{\partial V}{\partial \tau} - rV = 0
\end{aligned}
\tag{3.14}
$$

Multiplying by $\frac{\partial B}{\partial l}/\frac{\partial V}{\partial l}$ and subtracting from (3.9) yields:

$$\begin{aligned}
&\beta_1\left(\frac{\partial B}{\partial r}-\left(\frac{\partial B}{\partial l}\frac{\partial V}{\partial r}\right)\Big/\frac{\partial V}{\partial l}\right)+\frac{1}{2}\eta_1^2\left(\frac{\partial^2 B}{\partial r^2}-\left(\frac{\partial B}{\partial l}\frac{\partial^2 V}{\partial r^2}\right)\Big/\frac{\partial V}{\partial l}\right)\\
&\qquad+\frac{1}{2}\eta_2^2\left(\frac{\partial^2 B}{\partial l^2}-\left(\frac{\partial B}{\partial l}\frac{\partial^2 V}{\partial l^2}\right)\Big/\frac{\partial V}{\partial l}\right)\\
&+\rho\,\eta_1\eta_2\left(\frac{\partial B}{\partial r\partial l}-\left(\frac{\partial B}{\partial l}\frac{\partial V}{\partial r\partial l}\right)\Big/\frac{\partial V}{\partial l}\right)+\left(\frac{\partial B}{\partial l}\frac{\partial V}{\partial \tau}+rV\frac{\partial B}{\partial l}\right)\Big/\frac{\partial V}{\partial l}\\
&\qquad-\lambda_1\eta_1\left(\frac{\partial B}{\partial r}-\left(\frac{\partial B}{\partial l}\frac{\partial V}{\partial r}\right)\Big/\frac{\partial V}{\partial l}\right)-\frac{\partial B}{\partial \tau}-rB=0
\end{aligned}\tag{3.15}$$

We know:

$$\begin{aligned}
\frac{\partial V}{\partial r}=\frac{\partial^2 V}{\partial r^2} &= 0\\
\frac{\partial V}{\partial l} &= -l^{-2}\\
\frac{\partial^2 V}{\partial l^2} &= 2l^{-3}\\
\frac{\partial^2 V}{\partial r\partial l} &= 0\\
\frac{\partial V}{\partial \tau}=-\frac{\partial V}{\partial t} &= -1
\end{aligned}$$

and hence (3.15) simplifies to (3.13), the bond price equation independent of λ_2. By applying the bond pricing equation to one of the underlying variables, we have reduced the number of utility-dependent parameters. The resulting model has the same estimation complexity as a one factor model, but the dynamical explanatory power of a two-factor model.

3.2. Specific models

BS performed two empirical analyses ([**10**] and [**11**]) using the above general formulation of the two factor model. In both cases they assign specific functional forms to the drifts and volatilities of r and l. The free parameters in these functional forms are estimated by applying statistical methods to market data. Then, assuming the functional forms and parameter values to be the true values, they estimate the value of the market price of instantaneous interest rate risk, λ_1. To simplify the analysis, λ_1 is assumed to be an intertemporal constant.

In the first analysis [**10**], which makes use of Canadian bond data, BS use the assumption that the excess expected rate of return on the consol bond over the instantaneous rate of interest is proportional to the degree of long-term interest rate risk (as represented in (3.12)) to solve for the expected rate of return:

$$\beta_2(r,l,t) = l^2 - rl + \frac{\eta_2^2}{l} + \lambda_2\eta_2 \tag{3.16}$$

where λ_2 is assumed constant.

Additional assumptions are made to ensure that interest rates remain non-negative. The standard deviations of the instantaneous changes in interest rates are assumed proportional to their current levels:

$$\begin{aligned} \eta_1(r,l,t) &= r\sigma_1 \\ \eta_2(r,l,t) &= l\sigma_2 \\ \text{and } \beta_1(0,l,t) &\geq 0 \\ \text{which ensures } \beta_2(r,0,t) &\geq 0 \end{aligned} \tag{3.17}$$

where σ_1 and σ_2 are the constants of proportionality. To determine the functional form of $\beta_1(r,l,t)$, BS assume that the long-term interest rate l, contains information about future values of the instantaneous rate of interest, hence r regresses towards a function of l. That is, the assumption is made that:

$$d\ln r = \alpha\left(\ln l - \ln p - \ln r\right)dt + \sigma_1 dz_1 \tag{3.18}$$

From (3.1) we have:

$$\begin{aligned} d\ln r &= \left(\frac{\beta_1}{r} - \frac{\sigma_1^2}{2}\right)dt + \sigma_1 dz_1 \\ \Rightarrow \beta_1 &= r\left(\alpha\ln\left(\frac{l}{p\,r}\right) + \frac{\sigma_1^2}{2}\right) \end{aligned}$$

where α is the speed of reversion of $\ln r$ to $\ln\frac{l}{p}$, where p is a scaling factor. Equation (3.1) now becomes:

$$dr = r\left(\alpha\ln\left(\frac{l}{p\,r}\right) + \frac{\sigma_1^2}{2}\right)dt + r\sigma_1 dz_1 \tag{3.19}$$

Applying the assumptions (3.17) to equation (3.16):

$$\beta_2 = l\left(l - r + \sigma_2^2 + \sigma_2\lambda_2\right)$$

and hence (3.2) becomes:

$$dl = l\left(l - r + \sigma_2^2 + \sigma_2\lambda_2\right) + l\sigma_2\,dz_2 \tag{3.20}$$

The behaviour of the interest rates is now described by a system of non-linear stochastic differential equations, (3.19) and (3.20). To empirically estimate the values of the free parameters, that is p, α, σ_1, σ_2 and λ_2, BS linearise and discretise the system by approximating r and l with functions of $\ln r$ and $\ln l$.

In the second analysis [**11**], BS use US bond data for the period 1958–1979. The joint stochastic process for the two interest rates is assumed to have the following functional form:

$$dr = (a_1 + b_1(l - r))\, dt + r\sigma_1\, dz_1 \tag{3.21}$$

$$dl = l\,(a_2 + b_2 r + c_2 l)\, dt + l\sigma_2\, dz_2 \tag{3.22}$$

Again, the unanticipated changes in the interest rates are assumed to be proportional to their current levels, and the instantaneous rate is assumed to regress to the long-term interest rate, hence $b_1 > 0$. Allowing $a_1 < 0$, introduces the possibility of negative interest rates. BS acknowledge this flaw, but retain it due to its empirical tractability. They consider the resulting bond pricing model to be more significant than the properties of the linear approximation of the true stochastic interest rate process.

To obtain the drift term for the long-term interest rate, equation (3.12) was again solved for β_2 with the additional assumption that λ_2 is a linear function of the two interest rates, r and l. Therefore, as outlined in [**45**]:

$$\lambda_2 = k_0 + k_1 r + k_2 l \tag{3.23}$$

where k_0, k_1 and k_2 are constants. Solving (3.12) for β_2 and applying assumptions (3.17) and (3.23):

$$\begin{aligned}
\beta_2(r, l, t) &= k_0 l\sigma_2 + k_1 r l\sigma_2 + k_2\sigma_2 l^2 + \sigma_2^2 l + l^2 - rl \\
&= l\left(\left(k_0\sigma_2 + \sigma_2^2\right) + (k_1\sigma_2 - 1)\, r + (k_2\sigma_2 + 1)\, l\right) \\
&= l\,(a_2 + b_2 r + c_2 l)
\end{aligned}$$

The system of stochastic equations (3.21) and (3.22) is then discretised so that the unknown parameters: a_1, b_1, σ_1, a_2, b_2, c_2, σ_2 and ρ may be estimated from market data. Assuming these estimates are the true parameter values, λ_1 is estimated. λ_1 is assumed to be an intertemporal constant, an unrealistic assumption given the long time span of the market data.

The results of the empirical tests performed by BS are rather inconclusive since the inconsistencies observed can be attributed to one of four causes: model misspecification in the form of an omission of state variables, misspecification of the functional forms of the stochastic processes for r and l, market inefficiencies or measurement error. Attempts by BS to isolate the specific causes of the errors are again inconclusive.

Rebonato [**45**] examines the behaviour of this specific model formulation. He varies the parameters over a range of reasonable values and finds that the dynamics of the coupled system (3.21) and (3.22) tend to be quite unstable.

The long-term interest rate has a non-negligible probability of reaching very high values in a finite period of time. This instability is due to the specific functional form and parameterisation chosen by BS in the implementation of their model.

3.3. Conclusion

The specific model formulations and empirical investigations undertaken by BS admit criticism due to their lack of stability and inconclusiveness of results. This should, however, not influence the significance of their general model formulation, which is mathematically and economically sound. The general model is an equilibrium model; hence the drifts are real world drifts and internal consistency requirements needed to eliminate one of the market prices of risk are encapsulated in the choice of state variables. BS provide a methodology which is successful in reducing the complexity of modelling the term structure. An alternative choice for the specific functional form of each stochastic process can produce reasonable descriptions of the evolution of the term structure.

CHAPTER 4

Longstaff and Schwartz: A Two-Factor Equilibrium Model

Longstaff and Schwartz (LS) [**38**] developed a two-factor model of the term structure based on the framework of Cox, Ingersoll and Ross [**18**] discussed in Chapter 2. The two factors are the short-term interest rate and the instantaneous variance of changes in this rate (volatility of the short-term interest rate). Therefore the prices of contingent claims reflect the current levels of the interest rate and its volatility. The choice of interest rate volatility as the second state variable is supported by the fact that volatility is a key variable in contingent claim pricing.

4.1. General framework

4.1.1. The underlying economy. The term structure is modelled within a continuous time economy where physical investment is performed by a single stochastic constant-returns-to-scale technology[1]. The single good produced by this technology is either consumed or reinvested in production. The returns realised on physical investment are described by the stochastic process:

$$\frac{dQ}{Q} = (\mu X + \theta Y)\,dt + \sigma\sqrt{Y}dz_1 \tag{4.1}$$

where μ, θ and σ are positive constants, X and Y are state variables and z_1 is a Wiener process. X is an economic factor driving expected returns but having no effect on the uncertainty of production, while Y affects expected returns and volatility. This implies that expected returns and volatility of production are not perfectly correlated. To ensure a non-negative risk-free rate of interest, $\theta > \sigma^2$. The state variables are modelled by:

[1]Constant-returns-to-scale implies that the percentage increase in output equals the same percentage increase in all the inputs required by the technology [**44**]. Increasing-returns-to-scale implies economies of scale where the percentage change in output exceeds the percentage increase in inputs; while diseconomies of scale refer to decreasing-returns-to-scale where the percentage increase in output is lower than percentage increase in inputs.

$$dX = (a - bX)dt + c\sqrt{X}dz_2 \tag{4.2}$$
$$dY = (d - eY)dt + f\sqrt{Y}dz_3 \tag{4.3}$$

where a, b, c, d, e and f are positive and z_2, z_3 are Wiener processes. Since X is uncorrelated to production uncertainty, we require z_2 to be uncorrelated with z_1 and z_3.

Markets are assumed to be perfectly continuous and competitive, with a fixed number of homogeneous investors having time-additive preferences of the form:

$$\mathbb{E}_t \left[\int_t^{\infty} \exp(-\rho s) \ln(C(s)) ds \right]$$

where C_s is the time s level of consumption, ρ the utility discount function and $\mathbb{E}_t$ [] the expectation taken at time t. Here, future consumption is discounted to time t by the factor ρ, which indicates the decreasing current utility of consumption in the future. However, if consumption is delayed, the investor is able to invest more and benefit from greater consumption in the future. Each investor wishes to maximise their utility subject to the budget constraint denoted by

$$dW = W\frac{dQ}{Q} - Cdt$$

where W is the investor's wealth and hence at any time t, the investor may either consume wealth (denoted by C) or invest the wealth in the single technology (4.1). The derived utility function obtained as the solution to the maximisation problem, has the form[2]:

$$J(W, X, Y, t) = \frac{\exp(-\rho t)}{\rho} \ln W + G(X, Y, t) \tag{4.4}$$

Optimal consumption and investment amounts are also determined as part of the maximisation problem. CIR (1985) [**18**] determined the associated optimal consumption function to take the form:

$$C^*(W, X, Y, t) = \frac{\rho W}{1 - \exp(-\rho(T - t))}$$

where $T \to \infty$. Hence determining the optimal consumption level to be ρW, the equilibrium wealth dynamics are derived to be:

$$dW = (\mu X + \theta Y - \rho) W dt + \sigma W \sqrt{Y} dz_1 \tag{4.5}$$

[2]This is the same form of the indirect/derived utility function as described in Chapter 2 §2.5

Equations (4.2), (4.3) and (4.5) form a joint Markovian process, the current values of which completely describe the state of the economy.

Now rescale the state variables such that $x = X/c^2$ and $y = Y/f^2$ and define $H(x, y, \tau)$ as the value of a contingent claim with maturity τ. The value of this contingent claim satisfies the fundamental partial differential equation [**17**, Theorem 3]:

$$\begin{aligned} &\frac{x}{2}H_{xx} + \frac{y}{2}H_{yy} + (\gamma - \delta x)H_x \\ &+\left(\eta - \xi y - \left(\frac{-J_{WW}}{J_W}\right)\operatorname{covar}(dW, dY)\right)H_y - rH = H_\tau \end{aligned} \tag{4.6}$$

where[3] $\gamma = a/c^2$, $\eta = d/f^2$, $\delta = b$, $\xi = e$, r is the instantaneous risk-free rate of interest and $\operatorname{covar}(dW, dY)$ is the instantaneous covariance of changes in wealth W with changes in state variable Y. The coefficient of H_y includes a utility-dependent term which represents the risk premium associated with the level of production uncertainty governed by Y. From (4.3), (4.4) and (4.5), we may show[4]:

$$\left(\frac{-J_{WW}}{J_W}\right)\operatorname{covar}(dW, dY) = \lambda y$$

Hence the risk premium (market price of risk) is proportional to y. The proportionality factor λ is constant. This form of the risk premium has been endogenously derived from the specification of the underlying economy and so ensures consistency with the equilibrium characteristics of the economy. The

[3]Under this transformation the processes (4.2) and (4.3) become:

$$\begin{aligned} dx &= (\gamma - \delta x)dt + \sqrt{x}dz_2 \\ dy &= (\eta - \xi y)dt + \sqrt{y}dz_3 \end{aligned}$$

[4]From (4.4) we have:

$$\begin{aligned} J_W &= \frac{1}{\rho W}\, e^{-\rho t} \\ J_{WW} &= -\frac{1}{\rho W^2}\, e^{-\rho t} \\ \Rightarrow -\frac{J_{WW}}{J_W} &= \frac{1}{W} \end{aligned}$$

Also from (4.3) and (4.5) we find:

$$\operatorname{covar}(dW, dY) = WY\sigma f\,\mathbb{E}[dz_1\, dz_3]$$

Hence, since σ, f and $\mathbb{E}[dz_1\, dz_3]$ are constant, we have:

$$\left(\frac{-J_{WW}}{J_W}\right)\operatorname{covar}(dW, dY) = \lambda y \qquad \text{for } \lambda \text{ a constant.}$$

same cannot be said of the no-arbitrage models where the form of the risk premium is exogenously defined.

Specifying the instantaneous risk-free rate of interest r, in terms of the state variables x and y, will allow us to solve (4.6) for the price of any contingent claim subject to its initial and terminal boundary conditions. Since these prices will be in terms of unobservable state variables, we make a transformation to express the contingent claim prices in terms of intuitive and readily estimated economic variables. These two economic factors are the short-term interest rate r, and variance of changes in this short-term interest rate V. Hence information about the current term structure other than its current level is incorporated. This approach has the potential to produce contingent claim valuations more consistent with actual market prices than one-factor models.

4.1.2. The observable economic variables.

4.1.2.1. *Definition.* The equilibrium risk-free rate of interest is derived as the expected return on production less variance of production returns [**17**, Theorem 1]. That is[5]:

$$r = \alpha x + \beta y \tag{4.7}$$

where $\alpha = \mu c^2$ and $\beta = (\theta - \sigma^2) f^2$.

4.1.2.2. *Stochastic processes.* Applying Ito's Lemma to (4.7) we obtain the stochastic processes describing the evolution of r:

$$\begin{aligned} dr &= \alpha\, dx + \beta\, dy \\ &= \alpha(\gamma - \delta x)\, dt + \beta(\eta - \xi y)\, dt + \alpha\sqrt{x}\, dz_2 + \beta\sqrt{y}\, dz_3 \end{aligned} \tag{4.8}$$

Taking expected values of the above we obtain the expression for the instantaneous variance V, of changes in the short-term interest rate as:

$$\begin{aligned} \mathbb{E}\left[dr\right] &= \alpha\left(\gamma - \delta x\right) dt + \beta\left(\eta - \xi y\right) dt \\ \Rightarrow \left(\, \mathbb{E}\left[dr\right]\right)^2 &= 0 \end{aligned}$$

and since z_2 and z_3 are uncorrelated we have:

$$\mathbb{E}\left[(dr)^2\right] = \alpha^2 x dt + \beta^2 y dt$$

[5]From (4.1) the expected return less variance of the production process is $\mu X + \theta Y - \sigma^2 Y$. So, using the transformed state variables:

$$\begin{aligned} r &= \mu c^2 x + \theta f^2 y - \sigma^2 f^2 y \\ &= \mu c^2 x + (\theta - \sigma^2) f^2 y \end{aligned}$$

Hence

$$V = \frac{\operatorname{var}(dr)}{dt} = \mathbb{E}\left[(dr - \mathbb{E}[dr])^2\right]$$
$$= \mathbb{E}\left[(dr)^2\right] - (\,\mathbb{E}\,[dr])^2$$
$$= \alpha^2 x + \beta^2 y \tag{4.9}$$

Here the notational dependence on dt is suppressed since it is implicit in the context of the definition. Now solving (4.7) and (4.9) as a system of simultaneous equations, x and y may be found as functions of r and V. Hence, assuming $\alpha \neq \beta$ we have:

$$x = \frac{\beta\, r - V}{\alpha(\beta - \alpha)} \tag{4.10}$$

$$y = \frac{V - \alpha\, r}{\beta(\beta - \alpha)} \tag{4.11}$$

Now substituting these values of x and y into (4.8), the stochastic process for r, we have:

$$dr = \mu_r dt + \sigma_{1,r} dz_2 + \sigma_{2,r} dz_3 \tag{4.12}$$

where

$$\mu_r = \alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha}\, r - \frac{\xi - \delta}{\beta - \alpha}\, V$$
$$\sigma_{1,r} = \alpha\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}$$
$$\sigma_{2,r} = \beta\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}$$

Similarly, from (4.9) we derive the process for V as:

$$dV = \mu_V dt + \sigma_{1,V} dz_2 + \sigma_{2,V} dz_3 \tag{4.13}$$

where

$$\mu_V = \alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta - \xi)}{\beta - \alpha}\, r - \frac{\beta\xi - \alpha\delta}{\beta - \alpha}\, V$$
$$\sigma_{1,V} = \alpha^2\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}$$
$$\sigma_{2,V} = \beta^2\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}$$

Since the stochastic evolution of r depends on V and vice versa, the two processes are interdependent, forming a joint Markovian process.

4.1.2.3. *Expected value.* The methodology used to calculate the expected value of the short-term interest rate in the CIR model is applied to evaluate the unconditional expected values of $r(t)$ and $V(t)$. Since both $r(t)$ and $V(t)$ are linear combinations of x and y, we first calculate the expectations of x and y. The stochastic processes for x and y are:

$$dx = (\gamma - \delta x)dt + \sqrt{x}\, dz_2 \tag{4.14}$$

$$dy = (\eta - \xi y)dt + \sqrt{y}\, dz_3 \tag{4.15}$$

First, consider the integral form of (4.14):

$$x(t) = x(0) + \int_0^t (\gamma - \delta x(u))\, du + \int_0^t \sqrt{x(u)}\, dz_2$$

Taking expectations, we have:

$$\begin{aligned} \mathbb{E}\left[x(t)\right] &= x(0) + \int_0^t (\gamma - \delta\, \mathbb{E}\left[x(u)\right])\, du \\ \Rightarrow \frac{d}{dt}\, \mathbb{E}\left[x(t)\right] &= \gamma - \delta\, \mathbb{E}\left[x(t)\right] \end{aligned}$$

which implies

$$\begin{aligned} \frac{d}{dt}\left(e^{\delta t}\, \mathbb{E}\left[x(t)\right]\right) &= e^{\delta t}\left(\delta\, \mathbb{E}\left[x(t)\right] + \frac{d}{dt}\, \mathbb{E}\left[x(t)\right]\right) \\ &= \gamma\, e^{\delta t} \end{aligned}$$

Integrating both sides yields:

$$e^{\delta t}\, \mathbb{E}\left[x(t)\right] = x(0) + \frac{\gamma}{\delta}\left(e^{\delta t} - 1\right)$$

and hence we solve for $\mathbb{E}\left[x(t)\right]$ as:

$$\mathbb{E}\left[x(t)\right] = \frac{\gamma}{\delta}\left(1 - e^{-\delta t}\right) + e^{-\delta t} x(0) \tag{4.16}$$

If $x(0) = \frac{\gamma}{\delta}$, the mean reversion level, then $\mathbb{E}\left[x(t)\right] = \frac{\gamma}{\delta}$ for all t. Alternatively, for $x(0) \neq \frac{\gamma}{\delta}$, the long-run mean of $x(t)$ is:

$$\lim_{t \to \infty} \mathbb{E}\left[x(t)\right] = \frac{\gamma}{\delta} \tag{4.17}$$

Similarly, the long-run mean of $y(t)$ is calculated as:

$$\lim_{t \to \infty} \mathbb{E}\left[y(t)\right] = \frac{\eta}{\xi} \tag{4.18}$$

and hence:

$$\mathbb{E}[r] = \frac{\alpha\gamma}{\delta} + \frac{\beta\eta}{\xi} \tag{4.19}$$

$$\mathbb{E}[V] = \frac{\alpha^2\gamma}{\delta} + \frac{\beta^2\eta}{\xi} \tag{4.20}$$

4.1.2.4. *Variance.* Since z_2 and z_3 are uncorrelated, the variance of $r(t)$ (and similarly of $V(t)$) may be calculated as:

$$\operatorname{var}(r(t)) = \operatorname{var}(\alpha x + \beta y) = \alpha^2 \operatorname{var}(x) + \beta^2 \operatorname{var}(y) \tag{4.21}$$

where:

$$\operatorname{var}(x) = \mathbb{E}\left[x^2\right] - \mathbb{E}[x]^2 \tag{4.22}$$

In order to evaluate $\operatorname{var}(x)$ and hence $\operatorname{var}(r(t))$, we require the process $dx^2(t)$. Applying Ito's Lemma and making use of (4.14) we have:

$$\begin{aligned}
d\big(x^2(t)\big) &= 2x(t)\,dx(t) + dx(t)\,dx(t) \\
&= 2x(t)\left((\gamma - \delta x(t))dt + \sqrt{x(t)}\,dz_2\right) + x(t)dt \\
&= (2\gamma+1)\,x(t)dt - 2\delta\,x^2(t)\,dt + 2x^{\frac{3}{2}}(t)dz_2 \\
\Rightarrow x^2(t) &= x^2(0) + (2\gamma+1)\int_0^t x(u)du - 2\delta\int_0^t x^2(u)du + 2\int_0^t x^{\frac{3}{2}}(u)dz_2
\end{aligned}$$

Hence:

$$\mathbb{E}\left[x^2(t)\right] = x^2(0) + (2\gamma+1)\int_0^t \mathbb{E}[x(u)]\,du - 2\delta\int_0^t \mathbb{E}\left[x^2(u)\right]du$$

and

$$\frac{d}{dt}\,\mathbb{E}\left[x^2(t)\right] = (2\gamma+1)\ \mathbb{E}[x(t)] - 2\delta\,\mathbb{E}\left[x^2(t)\right]$$

Therefore:

$$\begin{aligned}
\frac{d}{dt}\left(e^{2\delta t}\,\mathbb{E}\left[x^2(t)\right]\right) &= e^{2\delta t}\left(2\delta\ \mathbb{E}\left[x^2(t)\right] + \frac{d}{dt}\,\mathbb{E}\left[x^2(t)\right]\right) \\
&= e^{2\delta t}\,(2\gamma+1)\ \mathbb{E}[x(t)]
\end{aligned}$$

Making use of (4.16), solve for $\mathbb{E}\left[x^2(t)\right]$ by integrating the above:

$$\begin{aligned} e^{2\delta t}\,\mathbb{E}\left[x^2(t)\right] &= x^2(0) + \left(\frac{\gamma^2}{\delta^2} + \frac{\gamma}{2\delta^2}\right)\left(1 - e^{\delta t}\right)^2 \\ &\quad + \left(\frac{2\gamma}{\delta} + \frac{1}{\delta}\right) x(0)\left(e^{\delta t} - 1\right) \\ \Rightarrow \mathbb{E}\left[x^2(t)\right] &= x^2(0)e^{-2\delta t} + \left(\frac{\gamma^2}{\delta^2} + \frac{\gamma}{2\delta^2}\right)\left(1 - e^{-\delta t}\right)^2 \\ &\quad + \left(\frac{2\gamma}{\delta} + \frac{1}{\delta}\right) x(0)e^{-\delta t}\left(1 - e^{-\delta t}\right) \end{aligned}$$

Now making use of identity (4.22) we calculate the variance of $x(t)$ to be:

$$\operatorname{var}(x) = \frac{\gamma}{2\delta^2}\left(1 - e^{-\delta t}\right)^2 + \frac{x(0)}{\delta}\,e^{-\delta t}\left(1 - e^{-\delta t}\right)$$

and the long run variance becomes:

$$\lim_{t\to\infty} \operatorname{var}(x) = \frac{\gamma}{2\delta^2}$$

Similarly, the long run variance of $y(t)$ is:

$$\lim_{t\to\infty} \operatorname{var}(y) = \frac{\eta}{2\xi^2}$$

Finally, by identity (4.21), the long run variance of $r(t)$ is:

$$\operatorname{var}(r) = \frac{\alpha^2\gamma}{2\delta^2} + \frac{\beta^2\eta}{2\xi^2} \tag{4.23}$$

Similarly, the long run variance of $V(t)$ may be calculated to be:

$$\operatorname{var}(V) = \frac{\alpha^4\gamma}{2\delta^2} + \frac{\beta^4\eta}{2\xi^2} \tag{4.24}$$

4.2. Equilibrium term structure

Let $P(r, V, \tau)$ be the price of a risk-free discount bond with time τ to maturity. The value of this discount bond must satisfy the fundamental partial differential equation for the value of a contingent claim (4.6),within the equilibrium economy outlined earlier.

4.2.1. Discount bond price. The equilibrium value of $P(x, y, \tau)$ may be found by solving (4.6) subject to the terminal boundary condition $P(x, y, 0) = 1$. This equilibrium value has the form:

$$P(x,y,\tau) = A^{2\gamma}(\tau)B^{2\eta}(\tau) \times \exp\Big(\kappa\tau + (\delta-\phi)\big(1-A(\tau)\big)x + (\nu-\psi)\big(1-B(\tau)\big)y\Big) \tag{4.25}$$

where

$$A(\tau) = \frac{2\phi}{(\delta+\phi)(e^{\phi\tau}-1)+2\phi} \tag{4.26}$$

$$B(\tau) = \frac{2\psi}{(\nu+\psi)(e^{\psi\tau}-1)+2\psi} \tag{4.27}$$

and

$$\begin{aligned} \nu &= \xi+\lambda \\ \phi &= \sqrt{2\alpha+\delta^2} \\ \psi &= \sqrt{2\beta+\nu^2} \\ \kappa &= \gamma(\delta+\phi)+\eta(\nu+\psi) \end{aligned}$$

Applying the change of variables from x and y to r and V, the discount bond price may be represented as[6]:

$$P(r,V,\tau) = A^{2\gamma}(\tau)B^{2\eta}(\tau)\exp\Big(\kappa\tau + C(\tau)r + D(\tau)V\Big) \tag{4.28}$$

---

[6]First, substituting (4.10) and (4.11) into the exponent in (4.25) we have:

$$\begin{aligned} &\kappa\tau + (\delta-\phi)(1-A(\tau))x + (\nu-\psi)(1-B(\tau))y \\ &= \kappa\tau + (\delta-\phi)(1-A(\tau))\frac{\beta r - V}{\alpha(\beta-\alpha)} + (\nu-\psi)(1-B(\tau))\frac{V-\alpha r}{\beta(\beta-\alpha)} \\ &= \kappa\tau + \left(\frac{\beta(\delta-\phi)(1-A(\tau))}{\alpha(\beta-\alpha)} - \frac{\alpha(\nu-\psi)(1-B(\tau))}{\beta(\beta-\alpha)}\right) r \\ &\quad + \left(\frac{(\nu-\psi)(1-B(\tau))}{\beta(\beta-\alpha)} - \frac{(\delta-\phi)(1-A(\tau))}{\alpha(\beta-\alpha)}\right) V \end{aligned}$$

From (4.26) and (4.27) we have:

$$\begin{aligned} 1-A(\tau) &= \frac{(\delta+\phi)(e^{\phi\tau}-1)A(\tau)}{2\phi} \\ 1-B(\tau) &= \frac{(\nu+\psi)(e^{\psi\tau}-1)B(\tau)}{2\psi} \end{aligned}$$

as well as

$$\begin{aligned} \delta^2-\phi^2 &= -2\alpha \\ \nu^2-\psi^2 &= -2\beta \end{aligned}$$

and so we may continue as

where

$$
\begin{aligned}
C(\tau) &= \frac{\alpha\phi(e^{\psi\tau}-1)B(\tau)-\beta\psi(e^{\phi\tau}-1)A(\tau)}{\phi\psi(\beta-\alpha)} \\
D(\tau) &= \frac{\psi(e^{\phi\tau}-1)A(\tau)-\phi(e^{\psi\tau}-1)B(\tau)}{\phi\psi(\beta-\alpha)}
\end{aligned}
$$

The utility-dependent market price of risk parameter λ, enters the bond price equation via the parameter ν. This means it need not be estimated separately, but only in a functional form with parameter ξ. The exponential form of the bond price solution (4.28) ensures the tractability of the closed-form solution and simplifies its calculation.

4.2.2. Yield to maturity. Let $Y(\tau)$ be the yield to maturity on a discount bond with term τ to maturity. Hence:

$$
\begin{aligned}
P(r,V,\tau) &= e^{-Y(\tau)\tau} \\
\Rightarrow e^{-Y(\tau)\tau} &= A^{2\gamma}(\tau)B^{2\eta}(\tau)\exp\left(\kappa\tau+C(\tau)r+D(\tau)V\right) \\
\Rightarrow \quad Y(\tau) &= -\frac{2\gamma\ln A(\tau)+2\eta\ln B(\tau)+\kappa\tau+C(\tau)r+D(\tau)V}{\tau} \qquad (4.29)
\end{aligned}
$$

which indicates that the yield on any discount bond with maturity τ, is a linear function of state variables r and V. We can show[7] that $\lim_{\tau\to 0} Y(\tau) = r$, which is consistent with our definition of r as the instantaneous short-term interest rate.

$$
\begin{aligned}
&\left(\frac{\beta(\delta-\phi)(1-A(\tau))}{\alpha(\beta-\alpha)}-\frac{\alpha(\nu-\psi)(1-B(\tau))}{\beta(\beta-\alpha)}\right)r \\
&+\left(\frac{(\nu-\psi)(1-B(\tau))}{\beta(\beta-\alpha)}-\frac{(\delta-\phi)(1-A(\tau))}{\alpha(\beta-\alpha)}\right)V \\
=&\left(-\frac{\beta(e^{\phi\tau}-1)A(\tau)}{\phi(\beta-\alpha)}+\frac{\alpha(e^{\psi\tau}-1)B(\tau)}{\psi(\beta-\alpha)}\right)r+\left(-\frac{(e^{\psi\tau}-1)B(\tau)}{\psi(\beta-\alpha)}+\frac{(e^{\phi\tau}-1)A(\tau)}{\phi(\beta-\alpha)}\right)V \\
=&\frac{\alpha\phi(e^{\psi\tau}-1)B(\tau)-\beta\psi(e^{\phi\tau}-1)A(\tau)}{\phi\psi(\beta-\alpha)}r+\frac{\psi(e^{\phi\tau}-1)A(\tau)-\phi(e^{\psi\tau}-1)B(\tau)}{\phi\psi(\beta-\alpha)}V
\end{aligned}
$$

[7]It is easy to see that both the numerator and denominator of $Y(\tau)$ tend to zero as τ tends to zero. Hence apply L'Hopital's Rule. We have:

$$
\begin{aligned}
\lim_{\tau\to 0} Y(\tau) &= \lim_{\tau\to 0} -\left(\frac{2\gamma A'(\tau)}{A(\tau)}+\frac{2\eta B'(\tau)}{B(\tau)}+\kappa+C'(\tau)r+D'(\tau)V\right) \\
&= -\left(-\gamma(\delta+\phi)-\eta(\nu+\psi)+\kappa-r\right) \\
&= r
\end{aligned}
$$

since, by definition $\kappa=\gamma(\delta+\phi)+\eta(\nu+\psi)$.

We may also show[8] that $\lim_{\tau\to\infty} Y(\tau) = \gamma(\phi - \delta) + \eta(\psi - \nu)$ which is a constant, independent of current values of r and V. This is consistent with the above analysis that r and V have long-run stationary distributions and hence the influence of the current level of the interest rate diminishes for yields far into the future.

4.2.3. Shape of the yield curve. One of the shortcomings of one factor yield curve models is the restrictions they place on possible shapes the term structure can assume. The advantage of the current model is that the dependence of discount bond prices on both the short-term interest rate and volatility of the short-term interest rate introduces greater freedom to the shape of the term structure. The sign of[9] $P_r(r, V, \tau)$ is indeterminate and so changes in the short-term interest rate may introduce varying, often opposite, changes on the yield curve at different maturities. Short-term interest rate volatility V may change while r remains constant. Changes in V may affect the slope and volatility of the yield curve, often to different degrees at various maturities, changing the overall shape of the term structure. In fact, very complex yield curve shapes can be obtained with relatively easy manipulations of the parameters of the underlying state variables r and V.

4.2.4. Term structure of volatility. For the purposes of option pricing, it is important to correctly fit the yield curve and the term structure of volatilities [**45**]. Determining the volatility at various points on the yield curve is equivalent to determining bond yield volatilities for various maturities τ. To calculate the yield volatility, first use Ito's Lemma to calculate bond price volatility and then use Ito's Lemma again to find the yield volatility.

Volatilities of discount bond prices for various maturities may be derived by applying Ito's Lemma to (4.28):

$$dP = -\frac{\partial P}{\partial \tau}\,dt + \frac{\partial P}{\partial r}\,dr + \tfrac{1}{2}\frac{\partial^2 P}{\partial r^2}\,dr dr + \frac{\partial P}{\partial V}\,dV + \tfrac{1}{2}\frac{\partial^2 P}{\partial V^2}\,dV dV$$

---

[8] Again, apply L'Hopital's Rule to $Y(\tau)$:

$$\begin{aligned}\lim_{\tau\to\infty} Y(\tau) &= \lim_{\tau\to\infty} -\left(\frac{2\gamma A'(\tau)}{A(\tau)} + \frac{2\eta'(\tau)}{B(\tau)} + \kappa + C'(\tau)r + D'(\tau)V\right)\\ &= -(-2\gamma\phi - 2\eta\psi + \kappa)\\ &= \gamma(\phi - \delta) + \eta(\psi - \nu)\end{aligned}$$

by the definition of κ.

[9] Subscript indicates a partial derivative.

$$\begin{aligned}\frac{dP}{P} &= \Big(-P_\tau P + C(\tau)\mu_r + D(\tau)\mu_V\Big)dt \\ &\quad + \Big(\tfrac{1}{2}C^2(\tau)(\sigma_{1,r}^2 + \sigma_{2,r}^2) + \tfrac{1}{2}D^2(\tau)(\sigma_{1,V}^2 + \sigma_{2,V}^2)\Big)dt \\ &\quad + \Big(C(\tau)\sigma_{1,r} + D(\tau)\sigma_{1,V}\Big)dz_2 + \Big(C(\tau)\sigma_{2,r} + D(\tau)\sigma_{2,V}\Big)dz_3\end{aligned}$$

and since z_2 and z_3 are uncorrelated the variance of bond price returns becomes:

$$\begin{aligned}\operatorname{var}\left(\frac{dP}{P}\right) &= \Big(C^2(\tau)(\sigma_{1,r}^2 + \sigma_{2,r}^2) + D^2(\tau)(\sigma_{1,V}^2 + \sigma_{2,V}^2) \\ &\qquad +2C(\tau)D(\tau)(\sigma_{1,r}\sigma_{1,V} + \sigma_{2,r}\sigma_{2,V})\Big)dt\end{aligned} \tag{4.30}$$

From (4.28) we have:

$$\begin{aligned}C^2(\tau) &= \frac{\alpha^2\phi^2(e^{\psi\tau}-1)^2B^2(\tau) + \beta^2\psi^2(e^{\phi\tau}-1)^2A^2(\tau)}{\phi^2\psi^2(\beta-\alpha)^2} \\ &\quad - \frac{2\alpha\beta\psi\phi(e^{\psi\tau}-1)(e^{\phi\tau}-1)A(\tau)B(\tau)}{\phi^2\psi^2(\beta-\alpha)^2} \\ D^2(\tau) &= \frac{\psi^2(e^{\phi\tau}-1)^2A^2(\tau) + \phi^2(e^{\psi\tau}-1)^2B^2(\tau)}{\phi^2\psi^2(\beta-\alpha)^2} \\ &\quad - \frac{2\psi\phi(e^{\psi\tau}-1)(e^{\phi\tau}-1)A(\tau)B(\tau)}{\phi^2\psi^2(\beta-\alpha)^2} \\ C(\tau)D(\tau) &= \frac{(\alpha+\beta)\psi\phi(e^{\psi\tau}-1)(e^{\phi\tau}-1)A(\tau)B(\tau)}{\phi^2\psi^2(\beta-\alpha)^2} \\ &\quad - \frac{\beta\psi^2(e^{\phi\tau}-1)^2A^2(\tau) + \alpha\phi^2(e^{\psi\tau}-1)^2B^2(\tau)}{\phi^2\psi^2(\beta-\alpha)^2}\end{aligned}$$

Also:

$$\begin{aligned}\sigma_{1,r}^2 + \sigma_{2,r}^2 &= V \\ \sigma_{1,V}^2 + \sigma_{2,V}^2 &= -\alpha\beta(\beta+\alpha)r + (\beta^2 + \alpha\beta + \alpha^2)V \\ \sigma_{1,r}\sigma_{1,V} + \sigma_{2,r}\sigma_{2,V} &= -\alpha\beta r + (\beta+\alpha)V\end{aligned}$$

Hence the variance, (4.30) may be expressed as:

$$\begin{aligned}&\operatorname{var}\left(\frac{dP}{P}\right) \\ &= \Big(C^2(\tau)V + D^2(\tau)\left(-\alpha\beta(\beta+\alpha)r + (\beta^2+\alpha\beta+\alpha^2)V\right) \\ &\quad +2C(\tau)D(\tau)(-\alpha\beta r + (\beta+\alpha)V)\Big)dt\end{aligned} \tag{4.31}$$

Grouping and simplifying the coefficients of r and V, the variance of bond price returns becomes:

$$
\begin{aligned}
\text{var}\left(\frac{dP}{P}\right)\Big/ dt \;&=\; \frac{\sigma^2_{P(\tau)}}{P^2} \\
&= \left(\frac{\alpha\beta\psi^2(e^{\phi\tau}-1)^2A^2(\tau)-\alpha\beta\phi^2(e^{\psi\tau}-1)^2B^2(\tau)}{\phi^2\psi^2(\beta-\alpha)}\right) r \\
&\quad + \left(\frac{\beta\phi^2(e^{\psi\tau}-1)^2B^2(\tau)-\alpha\psi^2(e^{\phi\tau}-1)^2A^2(\tau)}{\phi^2\psi^2(\beta-\alpha)}\right) V
\end{aligned}
\tag{4.32}
$$

Finally, to derive the volatility of bond yields, apply Ito's Lemma to the expression $Y(\tau) = -\ln P(\tau)/\tau$ as in [**45**]:

$$
\begin{aligned}
dY &= \frac{\partial Y}{\partial t}\,dt + \frac{\partial Y}{\partial P}\,dP + \frac{1}{2}\frac{\partial^2 Y}{\partial P^2}\,dPdP \\
&= \frac{\partial Y}{\partial t}\,dt + \frac{1}{2}\frac{\partial^2 Y}{\partial P^2}\,\sigma^2_{P(\tau)}\,dt - \frac{1}{\tau P(\tau)}(\mu_{P(\tau)}dt + \sigma_{P(\tau)}dz) \\
&= \mu_{Y(\tau)}dt + \sigma_{Y(\tau)}dz
\end{aligned}
\tag{4.33}
$$

and so:

$$
\sigma_{Y(\tau)} \;=\; -\,\frac{1}{\tau P(\tau)}\,\sigma_{P(\tau)}
\tag{4.34}
$$

The volatility of bond price returns and hence bond yield volatility is a function of the term to maturity τ, as well as the underlying variables r and V. This implies a term structure of volatilities that may assume a wider variety of shapes than is possible for one factor models.

4.2.5. Correlation between rates of various maturities. One of the advantages of a two factor model over a one factor model, is that it allows for imperfect correlation between rates of various maturities. To determine this correlation, we calculate the correlation between yields on discount bonds of various maturities.

From (4.29), we may write the yield to maturity of a τ maturity bond as[10]:

$$
Y(\tau) = \frac{M(\tau) - C(\tau)r - D(\tau)V}{\tau}
\tag{4.35}
$$

[10]Here $M(\tau) = -\frac{1}{\tau}(2\gamma\ln A(\tau) + 2\eta\ln B(\tau) + \kappa\tau)$

Applying Ito's Lemma to (4.35) and making use of (4.12) and (4.13):

$$\begin{aligned} dY(\tau) &= \frac{\partial Y}{\partial t}\,dt + \frac{\partial Y}{\partial r}\,dr + \frac{\partial Y}{\partial V}\,dV + \frac{1}{2}\frac{\partial^2 Y}{\partial r^2}\,drdr + \frac{1}{2}\frac{\partial^2 Y}{\partial V^2}\,dVdV \\ &= \mu_{Y(\tau)}dt + \left(\frac{\partial Y}{\partial r}\,\sigma_{1,r} + \frac{\partial Y}{\partial V}\,\sigma_{1,V}\right)dz_2 + \left(\frac{\partial Y}{\partial r}\,\sigma_{2,r} + \frac{\partial Y}{\partial V}\,\sigma_{2,V}\right)dz_3 \\ &= \mu_{Y(\tau)}dt - \tfrac{1}{\tau}\Big(C(\tau)\sigma_{1,r} + D(\tau)\sigma_{1,V}\Big)dz_2 \\ &\quad -\tfrac{1}{\tau}\Big(C(\tau)\sigma_{2,r} + D(\tau)\sigma_{2,V}\Big)dz_3 \end{aligned}$$

and so:

$$\begin{aligned} dY(\tau_1)dY(\tau_2) &= \frac{1}{\tau_1\tau_2}\Big[\big(C(\tau_1)\sigma_{1,r} + D(\tau_1)\sigma_{1,V}\big)\big(C(\tau_2)\sigma_{1,r} + D(\tau_2)\sigma_{1,V}\big) \\ &\quad +\big(C(\tau_1)\sigma_{2,r} + D(\tau_1)\sigma_{2,V}\big)\big(C(\tau_2)\sigma_{2,r} + D(\tau_2)\sigma_{2,V}\big)\Big]dt \end{aligned} \tag{4.36}$$

By the definition of correlation we have:

$$\operatorname{corr}(dY(\tau_1), dY(\tau_2)) = \rho_{Y(\tau_1)Y(\tau_2)} = \frac{\operatorname{covar}(dY(\tau_1), dY(\tau_2))}{\sqrt{\operatorname{var}(dY(\tau_1))\ \operatorname{var}(dY(\tau_2))}} \tag{4.37}$$

Since $Y(\tau_1)$ and $Y(\tau_2)$ are Wiener processes, $\mathbb{E}[dY(\tau_1)] = \mathbb{E}[dY(\tau_1)] = 0$ and hence:

$$\operatorname{covar}(dY(\tau_1), dY(\tau_2)) = \mathbb{E}[dY(\tau_1)\,dY(\tau_2)]$$

and (4.37) becomes:

$$\rho_{Y(\tau_1)Y(\tau_2)} = \frac{\mathbb{E}[dY(\tau_1)\,dY(\tau_2)]}{\sigma_{Y(\tau_1)}\sigma_{Y(\tau_2)}dt}$$

Finally, using (4.36) the correlation may be expressed as:

$$\begin{aligned} &\rho_{Y(\tau_1)Y(\tau_2)} \\ &= \frac{1}{\tau_1\tau_2\,\sigma_{Y(\tau_1)}\,\sigma_{Y(\tau_2)}}\Big[\big(C(\tau_1)\sigma_{1,r} + D(\tau_1)\sigma_{1,V}\big)\big(C(\tau_2)\sigma_{1,r} + D(\tau_2)\sigma_{1,V}\big) \\ &\quad +\big(C(\tau_1)\sigma_{2,r} + D(\tau_1)\sigma_{2,V}\big)\big(C(\tau_2)\sigma_{2,r} + D(\tau_2)\sigma_{2,V}\big)\Big] \end{aligned}$$

where the variances, $\sigma^2_{Y(\tau_1)}$ and $\sigma^2_{Y(\tau_2)}$ may be obtained by taking expectations of (4.35) or by making use of bond price volatilities and (4.34).

4.3. Option pricing

LS [**38**] derive an analytical formula for the value of a European option on a default-free discount bond. The option value is derived from within the equilibrium framework and is hence a function of the two factors determining discount bond prices. The explicit dependence on the current interest rate volatility is a desirable characteristic, since volatility is one of the key determinants of option values.

Let $C(r, V, \tau; K, T)$ be the price of a European call option with maturity τ, strike price K, on a discount bond with term to maturity $\tau + T$. Since the value of this call option must satisfy the fundamental pricing equation (4.6), LS show the closed-form solution to be:

$$\begin{aligned} C(r,V,\tau;K,T) &= P(r,V,\tau+T)\Psi(\theta_1,\theta_2;4\gamma,4\eta,\omega_1,\omega_2) \\ &\quad -KP(r,V,\tau)\Psi(\theta_3,\theta_4;4\gamma,4\eta,\omega_3,\omega_4) \end{aligned} \tag{4.38}$$

where $\Psi(\cdot)$ is the bivariate non-central chi-square distribution and $\theta_1, \theta_2, \theta_3, \theta_4$, $\omega_1, \omega_2, \omega_3$ and ω_4 are functions of the parameters of the underlying state variables. Since the two variates of each bivariate distribution are independent, the joint density is a product of the individual densities of the two variates, each being non-central chi-square[11].

Since the variance of the interest rate follows a stochastic process, this model allows option pricing with stochastic volatility. This is one of the few interest rate option models allowing for closed-form solutions with stochastic volatility [**45**].

To extend this methodology, LS [**39**] use a separation of variables technique to incorporate the initial observed term structure. This extended model incorporates actual discount functions to determine the initial option prices. This means that it cannot be used to price simple discount bonds. However, all other European interest rate contingent claims may be valued.

Consider the value of a European default-free contingent claim $H(x, y, \tau)$, with payoff at maturity $H(x, y, 0) = G(x, y)$. (The contingent claim is a function of the transformed, unobservable state variables x and y.) The value of this contingent claim may be factorised into its forward value and a discount function, with the same maturity date as the contingent claim:

$$H(x,y,\tau) = P(x,y,\tau)\, M(x,y,\tau) \tag{4.39}$$

where $P(x, y, \tau)$ is a discount bond and $M(x, y, \tau)$ the forward value of the contingent claim. Calculating all required partial derivatives we rewrite (4.6) as:

$$\begin{aligned} &M\Big(\tfrac{x}{2}P_{xx} + \tfrac{y}{2}P_{yy} + (\gamma-\delta x)P_x + (\eta-\nu y)P_y - (\alpha x+\beta y)P - P_\tau\Big) \\ &\quad +P\Big(\tfrac{x}{2}M_{xx} + \tfrac{y}{2}M_{yy} + (\gamma-\delta x)M_x + (\eta-\nu y)M_y - M_\tau\Big) \\ &\quad +xP_xM_x + yP_yM_y = 0 \end{aligned} \tag{4.40}$$

[11]For details of the non-central chi-square distribution, see Chapter 2 on the CIR model.

However, since $P(x,y,\tau)$ also satisfies (4.6), we have:

$$\tfrac{x}{2}P_{xx} + \tfrac{y}{2}P_{yy} + (\gamma - \delta x)P_x + (\eta - \nu y)P_y - (\alpha x + \beta y)P - P_\tau = 0$$

and (4.40) becomes:

$$\begin{aligned} P\Big(\tfrac{x}{2}M_{xx} + \tfrac{y}{2}M_{yy} + (\gamma - \delta x)M_x + (\eta - \nu y)M_y - M_\tau\Big) \\ +xP_xM_x + yP_yM_y = 0 \\ \Rightarrow \tfrac{x}{2}M_{xx} + \tfrac{y}{2}M_{yy} + \Big(\gamma - \delta x + x\frac{P_x}{P}\Big)M_x \\ +\Big(\eta - \nu y + y\frac{P_y}{P}\Big)M_y - M_\tau = 0 \end{aligned} \tag{4.41}$$

with terminal boundary condition $M(x,y,0) = G(x,y)$. The terms dependent on the price of the discount bond, P_x/P and P_y/P, can be evaluated from (4.25) as:

$$\begin{aligned} P_x &= P(\delta - \phi)(1 - A(\tau)) \\ P_y &= P(\nu - \psi)(1 - B(\tau)) \end{aligned}$$

and hence the forward value of a European contingent claim may be found by solving (4.41) subject to an appropriate terminal boundary condition. The present value of this contingent claim is found by discounting the forward price by the unit discount bond (with appropriate maturity) observable from the current term structure.

From (4.41) we may consider the forward value of the contingent claim $M(x,y,\tau)$, to be the expected value of the terminal payoff $G(x,y)$, hence:

$$M(x,y,\tau) = \mathbb{E}\left[G(x,y)\right]$$

where the expectation is taken over the joint probability distribution of x and y implied by the risk adjusted processes:

$$dx = \Big(\gamma - \delta x + \frac{P_x}{P}\,x\Big)dt + \sqrt{x}dz_2 \tag{4.42}$$

$$dy = \Big(\eta - \nu x + \frac{P_y}{P}\,y\Big)dt + \sqrt{y}dz_3 \tag{4.43}$$

The two processes above are square root processes, much like in the CIR model, and hence produce distributions related to the non-central chi-square distribution. Since x and y are independent, the joint distribution (the bivariate, non-central chi-square distribution) is a product of the two distributions.

The advantage of this extension to the model is that all information from the current term structure, as well as dynamics of the state variables are incorporated in the pricing. However, this approach loses the general equilibrium consistency, since not all discount bond prices are endogenously determined.

4.4. Conclusion

In developing a two factor model, LS overcome one of the most frequently cited criticisms of one factor models: the perfect correlation of instantaneous returns on bonds of all maturities. The model produces closed-form option prices for the case of stochastic volatility; this is a highly desirable feature few other models can produce. Additionally, the very flexible functional form of the model allows for very complicated shapes of the yield curve to be obtained with relative ease. However, this flexible functional form makes calibration rather difficult. The flexible functionality allows almost any market-observed term structure to be fitted, but this does not necessarily ensure meaningful term structure dynamics. One of the inevitable side effects of increasing the numbers of factors is the increased complexity; here, pricing of a simple European option requires evaluation of the bivariate non-central chi-square distribution.

CHAPTER 5

Langetieg's Multi-Factor Equilibrium Framework

The term structure of interest rates is embedded in the macro-economic system and is related to various economic factors. For this reason, Langetieg [**36**] proposes a model that can accommodate an arbitrary number of economic variables. The model is essentially an extension of Vasicek's term structure model [**50**], studied in Chapter 1, with multiple sources of uncertainty.

5.1. Underlying assumptions

Langetieg makes certain assumptions which allow for a mathematically tractable, intuitively sound model:

Assumption 1. The set of stochastic economic factors which are related to the interest rate term structure follow a joint elastic random walk.

Assumption 2. The instantaneous risk-free rate of interest may be expressed as a linear combination of these factors.

Assumption 3. The market prices of risk of the factors are deterministic, that is, they are either constants or function of time only.

The assumption of an elastic random walk means that the Vasicek model, which incorporates a univariate elastic random walk, is extended to a multivariate elastic random walk. Vasicek does not assume the functional form of the bond price, but derives it from the following assumptions (which apply to Langetieg's model as well):

- Bond prices are functionally related to certain stochastic factors.
- These underlying factors follow a specific stochastic process.
- The markets are sufficiently perfect to allow for a no arbitrage equilibrium to be reached.

5.2. Choice of generating process

There exists empirical evidence to support both the random walk and the elastic random walk as generating processes for stochastic factors within a macro-economic system. Therefore we may conclude that the generating process for the short-term interest rate is adequately described by:

$$dr(t) = \big(a(t) + b(t)r(t)\big)dt + \sigma(t)dz(t)$$

where $a(t)$, $b(t)$ and $\sigma(t)$ are either constants or functions of time. $a(t) + b(t)r(t)$ is the stochastic[1] instantaneous drift and $\sigma(t)$ the deterministic instantaneous variance of $r(t)$. The behaviour of $r(t)$ is determined by the value of $b(t)$ since, for:

- $b < 0$, r tends to $-\frac{a}{b}$,
- $b = 0$, the generating process for r simplifies to a random walk,
- $b > 0$, r explodes in finite time since it is repelled by the level $-\frac{a}{b}$.

Under the random walk generating process, short-term interest rates drift to positive and negative infinity with probability one. The elastic random walk with $b < 0$ eliminates this problem. It does, however, allow transient occurrences of negative interest rates and hence is not an appropriate model when short-term interest rates are close to zero. Negative interest rates are completely eliminated by setting the variance coefficient proportional to r^α, $\alpha > 0$. In the case of the Cox, Ingersoll and Ross model [**18**], $\alpha = \frac{1}{2}$ (see Chapter 2). This creates a natural reflecting barrier at $r = 0$, but introduces mathematical complexity which is difficult to implement in the multivariate case where the underlying factors are stochastic. Langetieg makes use of an elastic random walk process, with the assumption that the short-term interest rate is sufficiently above zero to make the probability of negative interest rates, in finite time, negligible.

5.3. Multivariate elastic random walk

From Assumption 1 the n underlying factors follow a multivariate joint elastic random walk[2]:

$$dx_i = \alpha_i(t,x)dt + \sigma_i(t,x)dz_i \quad i = 1,\ldots,n$$

where $\alpha_i(t,x) = a_i + B_{ij}\,x_i$. In matrix notation this linear system of equations becomes:

$$dx = (a + Bx)\,dt + \sigma\,dz \tag{5.1}$$

where

$$\begin{aligned} dx' &= [dx_1\,dx_2\,\ldots\,dx_n] \\ a' &= [a_1\,a_2\,\ldots\,a_n] \end{aligned}$$

---

[1] It is stochastic due to the functional dependence on $r(t)$.

[2] As usual dz_i is the standard Wiener process with:

- $\mathbb{E}[dz_i] = 0$,
- $\mathbb{E}[dz_i\,dz_i] = dt$.

$$
\begin{aligned}
B &= n \times n \text{ matrix, elements } B_{ij} \\
(\sigma\, dz)' &= [\sigma_1 dz_1\ \sigma_2 dz_2\ \ldots\ \sigma_n dz_n]
\end{aligned}
$$

The short-term interest rate is expressed as a linear combination of the underlying stochastic factors (Assumption 2), hence:

$$r = w_0 + \sum_{i=1}^{n} w_i x_i = w_0 + w' x \tag{5.2}$$

where

x – vector of stochastic factors characterising the underlying economic system,

w – vector of weights which are either constants or functions of time.

The solution to (5.1) has the form[3]:

$$
\begin{aligned}
x(t) = \psi(t - t_0) \biggl(x(t_0) &+ \int_{t_0}^{t} \psi(s - t_0)^{-1} a\, ds \\
&+ \int_{t_0}^{t} \psi(s - t_0)^{-1} \sigma\, dz(s) \biggr) \qquad t \geq t_0
\end{aligned} \tag{5.3}
$$

where $\psi(t - t_0)$ is the matrix solution [**33**], [**3**] to:

$$\frac{d\psi(t - t_0)}{dt} = B\psi(t - t_0) \qquad \text{with } \psi(t_0 - t_0) = I$$

In the special case where B is a constant:

$$\psi(t - t_0) = \exp\left(B(t - t_0)\right)$$

and (5.3) becomes

[3] The deterministic system of equations corresponding to (5.1) is:

$$dx = (a + Bx)\, dt \tag{5.4}$$

This is a linear system, which has a solution of the form

$$x(t) = \psi(t - t_0)\nu(t)$$

where $\nu(t)$ is some function of time and $\psi(t - t_0)$ is the solution to the homogeneous matrix equation

$$d\psi(t - t_0) = B\psi(t - t_0)dt$$

with initial condition $\psi(t_0 - t_0) = I$, and hence it is the fundamental matrix of the system (5.4).

Matrix equation (5.4) now becomes:

$$x(t) = \psi(t-t_0)x(t_0) + \int_{t_0}^{t} \psi(t-s)\,a\,ds + \int_{t_0}^{t} \psi(t-s)\,\sigma\,dz(s) \quad t \geq t_0 \tag{5.5}$$

For this special case[4], the expected value and covariance matrix of $x(t)$ and $x(t^*)$ (t and t^* are future points in time) are then calculated to be[5]:

$$\mathbb{E}_{t_0}[x(t)] = \psi(t-t_0)x(t_0) + \int_{t_0}^{t} \psi(t-s)\,a\,ds \tag{5.6}$$

$$\text{covar}_{t_0}(x(t), x(t^*)) = \int_{t_0}^{t\wedge t^*} \psi(t-s)\,\Sigma\,\psi(t^*-s)'ds \tag{5.7}$$

$$\begin{aligned}
\frac{\partial\psi(t-t_0)}{\partial t}\nu(t) + \psi(t-t_0)\frac{\partial\nu(t)}{\partial t} &= a + B\psi(t-t_0)\nu(t) \\
B\psi(t-t_0)\nu(t) + \psi(t-t_0)\frac{\partial\nu(t)}{\partial t} &= a + B\psi(t-t_0)\nu(t) \\
\psi(t-t_0)\frac{\partial\nu(t)}{\partial t} &= a \\
\Rightarrow \nu(t) &= \nu(t_0) + \int_{t_0}^{t} \psi(s-t_0)^{-1} a\,ds \\
\Rightarrow \nu(t) &= x(t_0) + \int_{t_0}^{t} \psi(s-t_0)^{-1} a\,ds
\end{aligned}$$

and hence the solution to the deterministic system is:

$$x(t) = \psi(t-t_0)\left(x(t_0) + \int_{t_0}^{t} \psi(s-t_0)^{-1} a\,ds\right)$$

Now consider the system of stochastic equations

$$d\nu = \psi(t-t_0)^{-1}(a\,dt + \sigma\,dz)$$

which has solution

$$\nu(t) = \nu(t_0) + \int_{t_0}^{t} \psi(s-t_0)^{-1} a\,ds + \int_{t_0}^{t} \psi(s-t_0)^{-1}\sigma\,dz$$

Applying Ito's Lemma to $x(t) = \psi(t-t_0)\nu(t)$

$$\begin{aligned}
dx &= d\psi(t-t_0)\,\nu + \psi(t-t_0)\,d\nu \\
&= B\psi(t-t_0)\,\nu\,dt + a\,dt + \sigma\,dz \\
&= (a + Bx)\,dt + \sigma\,dz
\end{aligned}$$

which is the original differential system (5.1). Hence, we have shown that the solution to this system is

$$x(t) = \psi(t-t_0)\left(x(t_0) + \int_{t_0}^{t} \psi(s-t_0)^{-1} a\,ds + \int_{t_0}^{t} \psi(s-t_0)^{-1}\sigma\,dz\right)$$

[4]The calculations are shown for the special case due to the simplified notation, however the more general case proceeds in a similar fashion.

where Σ is the covariance matrix:

$$\Sigma = \operatorname{covar}_{t_0}(\sigma dz, \sigma dz) = \mathbb{E}_{t_0}[\sigma dz \cdot \sigma dz']$$

with elements $\sigma_{ij} = \rho_{ij}\,\sigma_i\,\sigma_j$ where $\rho_{ij} = \mathbb{E}_{t_0}[dz_i \cdot dz_j]$.

5.4. The bond pricing model

Langetieg presents a general model for pricing default-free bonds. The model allows any number of underlying stochastic factors which follow an arbitrary Ito process. Although the form of the solution can always be obtained, an explicit solution can only be determined for certain distributions. The case considered here, where the underlying factors follow a multivariate elastic

---

[5] The covariance is calculated as follows:

$$\begin{aligned}
&\operatorname{covar}_{t_0}(x(t), x(t^*)) \\
&= \mathbb{E}_{t_0}\left[(x(t) - \mathbb{E}_{t_0}[x(t)])\,(x(t^*) - \mathbb{E}_{t_0}[x(t^*)])'\right] \\
&= \mathbb{E}_{t_0}\left[x(t)x(t^*)' - x(t)\,\mathbb{E}_{t_0}[x(t^*)]' - \mathbb{E}_{t_0}[x(t)]\,x(t^*)' + \mathbb{E}_{t_0}[x(t)]\,\mathbb{E}_{t_0}[x(t^*)]'\right] \\
&= \mathbb{E}_{t_0}\left[x(t)x(t^*)'\right] - \mathbb{E}_{t_0}[x(t)]\,\mathbb{E}_{t_0}[x(t^*)]'
\end{aligned}$$

Consider:

$$\begin{aligned}
&x(t)x(t^*)' \\
&= \psi(t-t_0)x(t_0)\,(\psi(t^*-t_0)x(t_0))' + \psi(t-t_0)x(t_0)\left(\int_{t_0}^{t^*} \psi(t^*-s)\,a\,ds\right)' \\
&\quad + \psi(t-t_0)x(t_0)\left(\int_{t_0}^{t^*} \psi(t^*-s)\,\sigma\,dz(s)\right)' + \int_{t_0}^{t} \psi(t-s)\,a\,ds\,(\psi(t^*-t_0)x(t_0))' \\
&\quad + \int_{t_0}^{t} \psi(t-s)\,\sigma\,dz(s)\,(\psi(t^*-t_0)x(t_0))' + \int_{t_0}^{t} \psi(t-s)\,\sigma\,dz(s)\left(\int_{t_0}^{t^*} \psi(t^*-s)\,\sigma\,dz(s)\right)'
\end{aligned}$$

and so:

$$\begin{aligned}
&\mathbb{E}_{t_0}\left[x(t)x(t^*)'\right] \\
&= \psi(t-t_0)x(t_0)\,(\psi(t^*-t_0)x(t_0))' + \psi(t-t_0)x(t_0)\left(\int_{t_0}^{t^*} \psi(t^*-s)\,a\,ds\right)' \\
&\quad + \int_{t_0}^{t} \psi(t-s)\,a\,ds\,(\psi(t^*-t_0)x(t_0))' + \int_{t_0}^{t\wedge t^*} \psi(t-s)\,\Sigma\,\psi(t^*-s)'ds
\end{aligned}$$

Also:

$$\begin{aligned}
&\mathbb{E}_{t_0}[x(t)]\,\mathbb{E}_{t_0}[x(t^*)]' \\
&= \psi(t-t_0)x(t_0)\,(\psi(t^*-t_0)x(t_0))' + \psi(t-t_0)x(t_0)\left(\int_{t_0}^{t^*} \psi(t^*-s)\,a\,ds\right)' \\
&\quad + \int_{t_0}^{t} \psi(t-s)\,a\,ds\,(\psi(t^*-t_0)x(t_0))'
\end{aligned}$$

and hence (5.7) follows.

random walk, is mathematically tractable and hence an explicit solution can be found.

5.4.1. The Equilibrium bond price equation. Let $P = P(t, T, x)$ be the time t price of a default-free discount bond paying \$1 at time T. Applying Ito's Lemma, the bond price generating process is:

$$\begin{aligned} dP &= \frac{\partial P}{\partial t}\,dt + \sum_{i=1}^{n}\frac{\partial P}{\partial x_i}\,dx_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 P}{\partial x_i \partial x_j}\,dx_i dx_j \\ &= \frac{\partial P}{\partial t}\,dt + \sum_{i=1}^{n}\frac{\partial P}{\partial x_i}\,\alpha_i dt + \sum_{i=1}^{n}\frac{\partial P}{\partial x_i}\,\sigma_i dz_i \\ &\quad + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 P}{\partial x_i \partial x_j}\,\sigma_i\sigma_j\rho_{ij}\,dt \\ \Rightarrow \frac{dP}{P} &= \mu_P dt + \sum_{i=1}^{n}\beta_P^i dz_i \end{aligned} \tag{5.8}$$

where

$$\mu_P = \frac{1}{P}\left(\frac{\partial P}{\partial t} + \sum_{i=1}^{n}\frac{\partial P}{\partial x_i}\,\alpha_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 P}{\partial x_i \partial x_j}\,\sigma_i\sigma_j\rho_{ij}\right) \tag{5.9}$$

$$\beta_P^i = \frac{1}{P}\frac{\partial P}{\partial x_i}\,\sigma_i \tag{5.10}$$

that is, μ_P is the expected rate of return and β_P^i the unanticipated rate of return due to unexpected changes in factor x_i. A no arbitrage argument may be applied to derive the partial differential equation for the bond price. Let $\bar{P}$ be a portfolio of $n+1$ bonds P_j, with maturities T_j and portfolio weights γ_j, such that $\sum_{j=1}^{n+1}\gamma_j = 1$. Then

$$\bar{P} = \sum_{j=1}^{n+1}\gamma_j P_j = \gamma' P$$

where

$$\begin{aligned} \gamma' &= [\gamma_1\,\gamma_2\,\ldots\,\gamma_{n+1}] \\ P' &= [P_1\,P_2\,\ldots\,P_{n+1}] \end{aligned}$$

Hence, by Ito's Lemma the stochastic process for the bond portfolio is:

$$\begin{aligned}
\frac{d\bar{P}}{\bar{P}} &= \sum_{j=1}^{n+1} \gamma_j \mu_{P_j} dt + \sum_{j=1}^{n+1} \gamma_j \left(\sum_{i=1}^{n} \beta_{P_j}^i dz_i \right) \\
&= \sum_{j=1}^{n+1} \gamma_j \mu_{P_j} dt + \sum_{i=1}^{n} \left(\sum_{j=1}^{n+1} \gamma_j \beta_{P_j}^i \right) dz_i \\
&= \gamma' \mu \, dt + \sum_{i=1}^{n} \gamma' \beta^i dz_i
\end{aligned}$$

where

$$\begin{aligned}
\mu' &= [\mu_{P_1} \, \mu_{P_2} \, \ldots \, \mu_{P_{n+1}}] \\
(\beta^i)' &= [\beta_{P_1}^i \, \beta_{P_2}^i \, \ldots \, \beta_{P_{n+1}}^i]
\end{aligned}$$

For the portfolio to be risk-free, its return must be deterministic and hence the coefficients of the stochastic terms must be zero. That is:

$$\sum_{j=1}^{n+1} \gamma_j \beta_{P_j}^i = \gamma' \beta^i = 0 \quad \text{for } i = 1, \ldots, n \tag{5.11}$$

To prevent profitable arbitrage, the return on the bond portfolio must be the risk-free rate of interest and hence:

$$\sum_{j=1}^{n+1} \gamma_j \, \mu_{P_j} = r$$

$$\Rightarrow \sum_{j=1}^{n+1} \gamma_j \, (\mu_{P_j} - r) = \gamma'(\mu - r) = 0 \tag{5.12}$$

The vector γ is an $(n+1)$ vector and by (5.11) it is orthogonal to the n vectors β^i. Additionally, (5.12) implies that γ is orthogonal to vector $(\mu - r)$. An $(n+1)$ vector can be orthogonal to at most n independent vectors in an $(n+1)$ space; hence $(\mu - r)$ must be a linear combination of the β^is, $i = 1, \ldots, n$, and so:

$$\mu - r = \lambda_1 \beta^1 + \lambda_2 \beta^2 + \cdots + \lambda_n \beta^n \tag{5.13}$$

where λ_i, $i = 1, \ldots, n$ are scalars which may depend on time and the underlying stochastic factors. (5.13) is a matrix equation with $(n + 1)$ elements, one for each of the bonds in portfolio $\bar{P}$. Since λ_i, $i = 1, \ldots, n$ are independent of the bond maturity, that is they are the same for every bond P_j in portfolio $\bar{P}$, (5.13) applies to any bond P:

$$\mu_P - r = \lambda_1 \beta_P^1 + \lambda_2 \beta_P^2 + \cdots + \lambda_n \beta_P^n \tag{5.14}$$

This equation is the equilibrium condition for the instantaneous expected rate of return on a bond where each λ_i, $i = 1, \ldots, n$ may be interpreted as the market risk premium associated with factor x_i. Substituting for μ_P and β_P^i from (5.9) and (5.10), we derive the equilibrium bond price equation:

$$\begin{aligned}
&\frac{\partial P}{\partial t} + \sum_{i=1}^n \frac{\partial P}{\partial x_i}\,\alpha_i + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 P}{\partial x_i \partial x_j}\,\sigma_i \sigma_j \rho_{ij} - rP \\
&\quad = \lambda_1 \frac{\partial P}{\partial x_1}\,\sigma_1 + \cdots + \lambda_n \frac{\partial P}{\partial x_n}\,\sigma_n = \sum_{i=1}^n \frac{\partial P}{\partial x_i}\,\sigma_i \lambda_i
\end{aligned}$$

$$\Rightarrow \sum_{i=1}^n \frac{\partial P}{\partial x_i}(\alpha_i - \sigma_i \lambda_i) + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 P}{\partial x_i \partial x_j}\,\sigma_i \sigma_j \rho_{ij} + \frac{\partial P}{\partial t} - rP = 0 \tag{5.15}$$

with boundary condition $P(T, T) = 1$.

5.4.2. Risk-neutral measure. The solution of (5.15) takes the form:

$$\begin{aligned}
P(t, T, x) &= \mathbb{E}_t \left[\exp\left(A(T)\right)\right] \\
A(T) &= -\int_t^T r(v)\,dv - \int_t^T \tfrac{1}{2}\lambda\sigma' \Sigma^{-1} \lambda\sigma dv \\
&\quad - \int_t^T \lambda\sigma' \Sigma^{-1} \sigma\, dz(v)
\end{aligned} \tag{5.16}$$

where

$$\begin{aligned}
\lambda\sigma' &= [\lambda_1\sigma_1\ \lambda_2\sigma_2\ \ldots\ \lambda_n\sigma_n] \\
\sigma\, dz' &= [\sigma_1 dz_1\ \sigma_2 dz_2\ \ldots\ \sigma_n dz_n] \\
\Sigma &= \text{the covariance matrix}
\end{aligned}$$

This valuation is obtained by transforming to a risk-neutral measure via the Girsanov Theorem[6] and applying the Feynman–Kac Theorem (see [**45**]). Consider (5.14), that is:

$$\mu_P - r = \sum_{i=1}^n \lambda_i \beta_P^i$$

Substituting into (5.8), the bond price generating process, we have:

[6] The Girsanov Theorem, [**43**, Theorem 8.6.4], states: Let $Y(t)$ be an Ito process:

$$\begin{aligned}\frac{dP}{P} &= \Big(r + \sum_{i=1}^{n} \lambda_i \beta_P^i\Big) dt + \sum_{i=1}^{n} \beta_P^i dz_i \\ &= rdt + \sum_{i=1}^{n} \beta_P^i (\lambda_i dt + dz_i) \\ &= rdt + \sum_{i=1}^{n} \beta_P^i d\tilde{z}_i\end{aligned}$$

where, by the Girsanov Theorem , $d\tilde{z}_i = \lambda_i dt + dz_i,\ i = 1, \ldots, n$ are Brownian motions under the equivalent martingale measure. This change of measure technique has taken us into the risk-neutral world where the bond price grows at the deterministic risk-free rate of interest $r(t)$, and some stochastic component driven by the transformed Brownian motions, $\tilde{z}_i,\ i = 1, \ldots, n$. Now we solve for the market risk premium vector.

Suppressing the notational dependence on P, we have:

$$\begin{aligned}\lambda' &= [\lambda_1\, \lambda_2\, \ldots\, \lambda_n] \\ \beta' &= [\beta^1\, \beta^2\, \ldots\, \beta^n] \\ \text{and } (\lambda\beta)' &= [\lambda_1\beta^1\, \lambda_2\beta^2\, \ldots\, \lambda_n\beta^n]\end{aligned}$$

$$dY(t) = \beta(t)dt + \theta(t)dB(t) \qquad t \leq T$$

where $B(t)$ is a Brownian Motion. Assume there exist functions $\mu(t)$ and $\alpha(t)$ such that

$$\theta(t)\mu(t) = \beta(t) - \alpha(t)$$

where $\mu(t)$ satisfies

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \mu^2(s)\, ds\right)\right] < \infty$$

Set

$$Z_t = \exp\left(-\int_0^t \mu(s)\, dB(s) - \frac{1}{2}\int_0^t \mu^2(s)\, ds\right)$$

and

$$d\tilde{P} = Z_T\, dP \qquad t \leq T$$

then

$$\tilde{B}(t) = \int_0^t \mu(s)\, ds + B(t)$$

is a Brownian Motion under $\tilde{P}$ and the Ito process for $Y(t)$ may be represented in terms of $\tilde{B}(t)$ as

$$dY(t) = \alpha(t)\, dt + \theta(t)\, d\tilde{B}(t)$$

In vector notation, (5.14) becomes:

$$\begin{aligned} \beta'\lambda &= \mu - r \\ \Rightarrow (\beta\beta')^{-1}\beta\,\beta'\lambda &= (\beta\beta')^{-1}\beta\,(\mu - r) \\ \lambda &= (\mu - r)(\beta\beta')^{-1}\beta \\ \text{and so} \qquad \lambda\,dz &= (\lambda\beta)'(\beta\beta')^{-1}\beta\,dz \end{aligned}$$

This expression for the market risk premium vector takes into account the correlations between various Brownian motions. It can be shown[7] that

$$(\lambda\beta)'(\beta\beta')^{-1}\beta\,dz \equiv \lambda\sigma'\Sigma^{-1}\sigma dz$$

where $\lambda\sigma'$ and Σ^{-1} are defined in (5.16). Also[8]

$$\begin{aligned} \lambda\sigma'\Sigma^{-1}\sigma(\lambda\sigma'\Sigma^{-1}\sigma)' &= \lambda\sigma'\Sigma^{-1}\Sigma\Sigma^{-1}\lambda\sigma \\ &= \lambda\sigma'\Sigma^{-1}\lambda\sigma \end{aligned}$$

Hence the bond pricing formula (5.16) follows by Girsanov's Theorem.

5.4.3. Bond price solution. A solution to (5.16) can only be found if the probability density function of $A(T)$ is known and hence the expectation can be evaluated. Alternatively, numerical procedures may be applied directly to partial differential equation (5.15) to obtain the bond price solution.

Consider the solution to (5.16) under the assumption of a multivariate elastic random walk. Since the instantaneous short-term interest rate may be

---

[7] Here the result is shown in two dimensions, the extension to higher dimensions, follows naturally. Since $\beta' = [\beta^1\beta^2]$, the covariance matrix is:

$$\beta\beta' = \begin{bmatrix} (\beta^1)^2 & \beta^{1,2} \\ \beta^{1,2} & (\beta^2)^2 \end{bmatrix}$$

where $\beta^{1,2} = \rho\beta^1\beta^2$. The inverse of this covariance matrix is:

$$(\beta\beta')^{-1} = \frac{1}{(\beta^1)^2(\beta^2)^2 - (\beta^{1,2})^2} \begin{bmatrix} (\beta^2)^2 & -\beta^{1,2} \\ -\beta^{1,2} & (\beta^1)^2 \end{bmatrix}$$

From (5.10) we have $\beta^i = \frac{1}{P}\frac{\partial P}{\partial x_i}\,\sigma_i = \frac{1}{P}P_i\,\sigma_i$ and since $\sigma_{1,2} = \rho\sigma_1\sigma_2$:

$$\begin{aligned} (\lambda\beta)'(\beta\beta')^{-1}\beta\,dz &= \frac{P^4}{P_1^2\,P_2^2}\,\frac{1}{\sigma_1^2\sigma_2^2 - (\sigma_{1,2})^2}\,\frac{P_1^2\,P_2^2}{P^4} \\ &\quad \times(\sigma_1^2\sigma_2^2\lambda_1\,dz_1 - \rho\sigma_1^2\sigma_2^2\lambda_1\,dz_2 - \rho\sigma_1^2\sigma_2^2\lambda_2\,dz_1 + \sigma_1^2\sigma_2^2\lambda_2\,dz_2) \\ &= \frac{1}{1-\rho^2}\,[(\lambda_1 - \rho\lambda_2)\,dz_1 + (\lambda_2 - \rho\lambda_1)\,dz_2] \\ &= \lambda\sigma'\Sigma^{-1}\sigma dz \end{aligned}$$

[8] Since Σ and so Σ^{-1} are symmetric matrices we know $(\Sigma^{-1})' = \Sigma^{-1}$.

expressed as a linear function of the underlying stochastic factors, we have from (5.2) and (5.3):

$$\begin{aligned} r(v) &= w_0 + w'x(v) \\ &= w_0 + w'\left(\psi(v-t)x(t) + \int_t^v \psi(v-s)a ds + \int_t^v \psi(v-s)\sigma\, dz(s)\right) \end{aligned} \tag{5.17}$$

Hence (5.16) becomes:

$$\begin{aligned} A(T) \;=\; & -\int_t^T \left(w_0 + w'\psi(v-t)x(t) + w'\int_t^v \psi(v-s)a ds \right. \\ & \left. + w'\int_t^v \psi(v-s)\sigma dz(s)\right) dv \\ & - \int_t^T \tfrac{1}{2}\lambda\sigma'\Sigma^{-1}\lambda\sigma dv - \int_t^T \lambda\sigma'\Sigma^{-1}\sigma dz(v) \end{aligned} \tag{5.18}$$

Since $A(T)$ is normally distributed[9], we may evaluate the bond price as[10]:

$$P(t,T,x) = \exp\left(\mathbb{E}_t\left[A(T)\right] + \tfrac{1}{2}\operatorname{var}_t\left(A(T)\right)\right) \tag{5.19}$$

---

[9] $A(T)$ is normally distributed since it depends on a linear combination of stochastic factors generated by an elastic random walk, and the market prices of risk $\lambda_i,\ i = 1,\dots,n$, are deterministic.

[10] Consider a random variable X, which is normally distributed, $X \sim \Phi(\alpha, \beta)$. The probability density function of X is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

and hence the expected value of any function of X, $u(x)$ is calculated as:

$$\mathbb{E}\left[u(X)\right] = \int_{-\infty}^{\infty} u(x)f(x)dx$$

Let $u(x) = e^x$, and noting that $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$, we have

$$\begin{aligned} \mathbb{E}\left[e^X\right] &= \int_{-\infty}^{\infty} e^x \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \int_{-\infty}^{\infty} (1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots)\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &\quad + \frac{1}{2}\int_{-\infty}^{\infty} \frac{x^2}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \frac{1}{6}\int_{-\infty}^{\infty} \frac{x^3}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \cdots \end{aligned}$$

Since $\mathbb{E}\left[X\right] = \alpha$:

$$\mathbb{E}\left[(X-\alpha)^2\right] = \mathbb{E}\left[X^2\right] - \alpha^2 = \beta^2 \quad \Rightarrow \quad \mathbb{E}\left[X^2\right] = \beta^2 + \alpha^2$$

where[11]

$$\begin{aligned}
&\mathbb{E}_t\left[A(T)\right]\\
&= -\int_t^T \left(w_0 + w'\psi(v-t)x(t) + w'\int_t^v \psi(v-s)a\,ds + \tfrac{1}{2}\lambda\sigma'\Sigma^{-1}\lambda\sigma \right) dv\\
&\mathrm{var}_t\left(A(T)\right)\\
&= \int_t^T \int_t^T \int_t^v w'\psi(v-s)\Sigma\psi(v-s)'w\,ds\,dv\,dv + \int_t^T \lambda\sigma'\Sigma^{-1}\lambda\sigma\,dv\\
&\quad + 2\int_t^T \int_t^v w'\psi(v-s)\lambda\sigma\,ds\,dv
\end{aligned}$$

An explicit bond price is obtained by evaluating the above integrals.

---

and

$$\begin{aligned}
\mathbb{E}\left[(X-\alpha)^3\right] &= \mathbb{E}\left[X^3\right] - 3\alpha\,\mathbb{E}\left[X^2\right] + 2\alpha^3\\
&= \mathbb{E}\left[X^3\right] - 3\alpha\beta^2 - \alpha^3\\
&= \kappa\\
\Rightarrow \mathbb{E}\left[X^3\right] &= \kappa + 3\alpha\beta^2 + \alpha^3
\end{aligned}$$

for some $\kappa \in \mathbb{R}$. Hence

$$\begin{aligned}
\mathbb{E}\left[e^X\right] &= 1 + \alpha + \tfrac{1}{2}(\alpha^2+\beta^2) + \tfrac{1}{6}(\kappa + 3\alpha\beta^2 + \alpha^3) + \cdots\\
&= 1 + (\alpha + \tfrac{1}{2}\beta^2) + \tfrac{1}{2}(\alpha + \tfrac{1}{2}\beta^2)^2 + \cdots\\
&= e^{\alpha + \frac{1}{2}\beta^2}
\end{aligned}$$

[11]To calculate $\mathrm{var}_t\left(A(T)\right)$, the variance of $A(T)$, consider:

$$\begin{aligned}
&\mathrm{covar}_t\left(A(T)_v, A(T)_{v^*}\right)\\
&= \mathbb{E}\left[\left(A(T)_v - \mathbb{E}\left[A(T)_v\right]\right)\left(A(T)_{v^*} - \mathbb{E}\left[A(T)_{v^*}\right]\right)\right]\\
&= \mathbb{E}\Bigg[\int_t^T \int_t^T \left(w'\int_t^v \psi(v-s)\sigma dz(s) \left(w'\int_t^{v^*} \psi(v^*-s)\sigma dz(s)\right)'\right) dv\,dv^*\\
&\quad + \int_t^T \int_t^v w'\psi(v-s)\sigma\,dz(s)\,dv \left(\int_t^T \lambda\sigma'\Sigma^{-1}\sigma\,dz(v^*)\right)'\\
&\quad + \int_t^T \lambda\sigma'\Sigma^{-1}\sigma\,dz(v) \left(\int_t^T \int_t^{v^*} w'\psi(v^*-s)\sigma\,dz(s)\,dv^*\right)'\\
&\quad + \int_t^T \lambda\sigma'\Sigma^{-1}\sigma\,dz(v) \left(\int_t^T \lambda\sigma'\Sigma^{-1}\sigma\,dz(v^*)\right)'\Bigg]\\
&= \int_t^T \int_t^T \int_t^{v\wedge v^*} w'\psi(v-s)\Sigma\,\psi(v^*-s)'w\,ds\,dv\,dv^* + \int_t^T \int_t^v w'\psi(v-s)\lambda\sigma\,ds\,dv\\
&\quad + \int_t^T \int_t^{v^*} w'\psi(v^*-s)\lambda\sigma ds\,dv^* + \int_t^T \lambda\sigma'\Sigma^{-1}\lambda\sigma\,dv
\end{aligned}$$

5.4.4. Alternative approach to determining the bond price. Alternatively, considering equation (5.8)

$$\frac{dP}{P} = \mu_P \, dt + \sum_{i=1}^{n} \beta_P^i \, dz_i$$

we see that bond risk, that is the unexpected changes in bond price, is related to the bond gradient vector where the gradient is taken with respect to the underlying stochastic factors. From (5.19), the bond gradient vector is:

$$\begin{aligned} \frac{\partial P(t,T)}{\partial x'} &= \exp\Big(\mathbb{E}_t\left[A(T)\right] + \tfrac{1}{2}\operatorname{var}_t\left(A(T)\right)\Big)\frac{\partial\,\mathbb{E}_t\left[A(T)\right]}{\partial x} \\ &= P(t,T)\Big(-\int_t^T w'\psi(v-t)dv\Big) \end{aligned}$$

Let $\nabla(t,T)$ denote the normalised gradient vector, that is:

$$\begin{aligned} \nabla(t,T)' &= \frac{1}{P(t,T)}\frac{\partial P(t,T)}{\partial x'} \\ &= -\int_t^T w'\psi(v-t)dv \end{aligned} \tag{5.20}$$

This defines the bond risk per unit bond and hence $\nabla(t,T)$ represents the bond risk vector. Next, calculate the expected instantaneous equilibrium return on the bond. From (5.13), derived by means of a no arbitrage argument and (5.10), the definition of β^i, the expected instantaneous equilibrium return of the bond is:

$$\begin{aligned} \mu\, dt &= \big(r(t) + \lambda_1\beta^1 + \cdots + \lambda_n\beta^n\big)\, dt \\ &= \left(r(t) + \frac{\lambda_1\sigma_1}{P(t,T)}\frac{\partial P(t,T)}{\partial x_1} + \cdots + \frac{\lambda_n\sigma_n}{P(t,T)}\frac{\partial P(t,T)}{\partial x_n}\right) dt \\ &= \left(r(t) + \nabla(t,T)'\sigma\lambda\right) dt \end{aligned}$$

where $\nabla(t,T)' = \frac{1}{P(t,T)}\left[\frac{\partial P(t,T)}{\partial x_1} \cdots \frac{\partial P(t,T)}{\partial x_n}\right]$ as defined in (5.20). Hence the bond price generating process (5.8) becomes:

Since $\operatorname{var}_t\left(A(T)\right) = \operatorname{covar}_t\left(A(T)_v, A(T)_v\right)$ we have:

$$\begin{aligned} \operatorname{var}_t\left(A(T)\right) &= \int_t^T\int_t^T\int_t^v w'\psi(v-s)\Sigma\,\psi(v-s)'w\, ds\, dv\, dv \\ &\quad +2\int_t^T\int_t^v w'\psi(v-s)\lambda\sigma\, ds\, dv + \int_t^T \lambda\sigma'\Sigma^{-1}\lambda\sigma\, dv \end{aligned}$$

$$\begin{aligned}\frac{dP(t,T)}{P(t,T)} &= \mu\, dt + \sum_{i=1}^{n} \beta^i\, dz_i \\ &= (r(t) + \nabla(t,T)'\sigma\lambda)\, dt + \nabla(t,T)'\sigma\, dz\end{aligned}$$

By Ito's Lemma we have:

$$\begin{aligned}d\ln P(t,T) &= \frac{\partial \ln P(t,T)}{\partial P}\, dP + \frac{1}{2}\frac{\partial^2 \ln P(t,T)}{\partial P^2}\, dPdP \\ &= \frac{1}{P}\, dP - \frac{1}{2}\frac{1}{P^2}P^2\left(\nabla(t,T)'\Sigma\nabla(t,T)\right) dt \\ &= (r(t) + \nabla(t,T)'\sigma\lambda)\, dt + \nabla(t,T)'\sigma\, dz \\ &\quad -\tfrac{1}{2}\nabla(t,T)'\Sigma\nabla(t,T)dt\end{aligned}$$

$$\begin{aligned}\Rightarrow \ln P(s,T) - \ln P(t,T) &= \int_t^s r(v) + \nabla(v,T)'\sigma\lambda - \tfrac{1}{2}\nabla(v,T)'\Sigma\nabla(v,T)\, dv \\ &\quad + \int_t^s \nabla(v,T)'\sigma\, dz(v)\end{aligned}$$

$$\begin{aligned}\Rightarrow P(s,T) &= P(t,T)\exp\left(\int_t^s r(v) + \nabla(v,T)'\sigma\lambda\right. \\ &\quad \left. -\tfrac{1}{2}\nabla(v,T)'\Sigma\nabla(v,T)dv + \int_t^s \nabla(v,T)'\sigma dz(v)\right)\end{aligned} \tag{5.21}$$

where $P(s,T)$ is a known value of the bond at some time $s \geq t$. From (5.6):

$$\mathbb{E}_t\left[x(v)\right] = \psi(v-t)x(t) + \int_t^v \psi(v-u)a\, du$$

which implies

$$\begin{aligned}& w_0 + w'\left(\psi(v-t)x(t) + \int_t^v \psi(v-u)adu\right) \\ &= w_0 + w'\,\mathbb{E}_t\left[x(v)\right] \\ &= \mathbb{E}_t\left[r(v)\right]\end{aligned}$$

by (5.17), and so

$$r(v) = \mathbb{E}_t\left[r(v)\right] + w'\int_t^v \psi(v-u)\sigma\, dz(u)$$

Therefore (5.21) becomes

$$
\begin{aligned}
P(s,T) = P(t,T)\exp\Bigg(&\int_t^s \mathbb{E}_t\left[r(v)\right]dv + \int_t^s L(v,t)\,dv \\
&+ \int_t^s w' \int_t^v \psi(v-u)\sigma\,dz(u)\,dv + \int_t^s \nabla(v,T)'\sigma\,dz(v)\Bigg)
\end{aligned}
\tag{5.22}
$$

where $L(v,t) = \nabla(v,T)'\sigma\lambda - \frac{1}{2}\nabla(v,T)'\Sigma\nabla(v,T)$.

Now making use of (5.20), we have:

$$
\begin{aligned}
&\int_t^s w' \int_t^v \psi(v-u)\sigma\,dz(u)\,dv \\
&= \int_t^s \int_u^s w'\psi(v-u)\sigma dv\,dz(u) \\
&= -\int_t^s \nabla(u,s)'\sigma\,dz(u)
\end{aligned}
$$

and

$$
\begin{aligned}
\nabla(v,T)' &= -\int_v^T w'\psi(m-v)dm \\
&= -\int_v^s w'\psi(m-v)dm - \int_s^T w'\psi(m-v)dm \\
&= \nabla(v,s)' - \int_s^T w'\psi(m-v)dm
\end{aligned}
$$

Hence (5.22) simplifies to

$$
\begin{aligned}
P(s,T) = P(t,T)\exp\Bigg(&\int_t^s \mathbb{E}_t\left[r(v)\right]dv + \int_t^s L(v,t)\,dv \\
&- \int_t^s \int_s^T w'\psi(m-v)\sigma\,dm\,dz(v)\Bigg)
\end{aligned}
\tag{5.23}
$$

Now since $\psi(m-v) = \exp B(m-v)$ for constant B, we have $\psi(m-v) = \exp\left(B(m-s) + B(s-v)\right)$ and

$$\begin{aligned}
&\int_t^s \int_s^T w'\psi(m-v)\sigma \, dm \, dz(v) \\
&= \int_t^s \int_s^T w' \exp B(m-s) \exp B(s-v) \, \sigma \, dm \, dz(v) \\
&= \int_t^s \exp B(s-v) \int_s^T w' \exp B(m-s) \, \sigma \, dm \, dz(v) \\
&= -\int_t^s \exp B(s-v) \, \sigma \, \nabla(s,T)' \, dz(v) \\
&= -\int_t^s \psi(s-v) \, \sigma \, \nabla(s,T)' \, dz(v)
\end{aligned}$$

Also, from (5.5) and (5.6) we have:

$$\begin{aligned}
x(s) &= \psi(s-t)x(t) + \int_t^s \psi(s-v) \, a \, dv + \int_t^s \psi(s-v) \, \sigma \, dz(v) \\
\mathbb{E}_t\left[x(s)\right] &= \psi(s-t)x(t) + \int_t^s \psi(s-v) \, a \, dv \\
\Rightarrow \int_t^s \psi(s-v) \, \sigma \, dz(v) &= x(s) - \mathbb{E}_t\left[x(s)\right]
\end{aligned}$$

Hence (5.23) becomes:

$$\begin{aligned}
P(s,T) &= P(t,T) \exp\bigg(\int_t^s \mathbb{E}_t\left[r(v)\right] dv + \int_t^s L(v,t) \, dv \\
&\qquad + \nabla(s,T)' \int_t^s \psi(s-v)\sigma \, dz(v) \bigg) \\
&= P(t,T) \exp\bigg(\int_t^s \mathbb{E}_t\left[r(v)\right] dv + \int_t^s L(v,t) \, dv \\
&\qquad + \nabla(s,T)' \Big(x(s) - \mathbb{E}_t\left[x(s)\right] \Big) \bigg)
\end{aligned} \tag{5.24}$$

By the definition of $P(t,T)$, we have the boundary condition $P(T,T) = 1$. Setting $s = T$ in (5.24) above and noting that $\nabla(T,T) = 0$, solve for $P(t,T)$ as:

$$\begin{aligned}
P(t,T) &= \exp\bigg(-\int_t^T \mathbb{E}_t\left[r(v)\right] dv - \int_t^T L(v,t) \, dv \\
&\qquad + \nabla(T,T)' \Big(x(s) - \mathbb{E}_t\left[x(s)\right] \Big) \bigg) \\
&= \exp\bigg(-\int_t^T \mathbb{E}_t\left[r(v)\right] dv - \int_t^T L(v,t) \, dv \bigg)
\end{aligned} \tag{5.25}$$

This equation provides a simple and intuitive representation of the bond price. Changes in bond price may be attributed to the expected changes in the risk-free rate of interest and another term dependent on the underlying economic factors.

By definition $P(t,T) = \exp(-R(t,T)(T-t))$, and so the term structure equation may then be determined as:

$$R(t,T) = \frac{1}{T-t}\left(\int_t^T \mathbb{E}_t\left[r(v)\right] dv + \int_t^T L(v,t)\, dv\right)$$

5.5. Conclusion

Langetieg's model allows the incorporation of an arbitrary number of economic factors into the bond pricing equation. The short-term interest rate is specified as a linear combination of the economic factors, which are assumed to follow a joint elastic random walk. Hence the bond price becomes a function of this linear combination of economic factors. The ability to find a closed-form solution for the bond price depends on the type of process assumed for the economic factors. The elastic random walk results in a normal distribution which makes such a solution possible.

This model makes a theoretical rather than a practical contribution to interest rate modelling. Its contribution lies in the theoretical framework it provides for the incorporation of multiple economic factors. The model parameters are left unspecified, so application of the model requires the specification of the number and type of economic factors determining the short-term interest rate, the estimation of the parameters of the joint elastic random walk process followed by the economic factors and the estimation of the associated market prices of risk. This is a rather challenging exercise giving rise to many estimation complexities.

CHAPTER 6

The Ball and Torous Model

Ball and Torous (BT) [**4**] propose an equilibrium methodology to value contingent claims on risk-free zero coupon bonds. The resulting closed-form valuation formula is independent of investor preferences and eliminates the need for numerical estimations of utility-dependent factors.

The underlying state variable is the risk-free zero coupon bond directly. Its price is assumed to follow a Brownian Bridge process, ensuring that it converges to the face value at maturity. Also, since this underlying state variable is a tradable security, a preference-free, closed-form valuation formula for European options may be derived.

6.1. Holding period returns

Let $P(t,m)$ be the time t price of a risk-free zero coupon bond with maturity m. $P(m,m) = 1$. Define $\xi(t,m)$ to be the t-period log return on the zero coupon bond:

$$\xi(t,m) = \ln P(t,m) - \ln P(0,m) = \ln \frac{P(t,m)}{P(0,m)} \tag{6.1}$$

Since

$$\xi(0,m) = \ln \frac{P(0,m)}{P(0,m)} = 0$$

and

$$\xi(m,m) = \ln \frac{P(m,m)}{P(0,m)} = -\ln P(0,m)$$

we see that $\xi(t,m)$ is constrained at $t = 0$ and $t = m$.

The yield-to-maturity $\mu(m)$, is defined as the continuously compounded rate of return (per unit time) earned if the bond is bought at time $t = 0$ and held until maturity $t = m$. Hence:

$$\begin{aligned} P(m,m) &= P(0,m)\exp(\mu(m)m) \qquad\qquad m \in [0,\infty) \\ \Rightarrow \mu(m) &= \frac{-\ln P(0,m)}{m} \end{aligned}$$

Therefore, if the investor commits to holding the bond until maturity he earns a log return of $\mu(m)t$ after t units of time, where $t \leq m$. However, if the bond is not held until maturity the return diverges from the deterministic yield-to-maturity and is stochastic. Let $\eta(t,m)$ be the excess log return earned on the risk-free bond over time t. Therefore:

$$\eta(t,m) = \xi(t,m) - \mu(m)t \qquad t \in [0,m] \tag{6.2}$$

As illustrated by (6.2), the bond return may be decomposed into a deterministic and a stochastic component. Holding the bond for time t, a deterministic return of $\mu(m)t$ is earned. Selling the bond prior to maturity introduces the stochastic component $\eta(t,m)$, which reflects the changing market conditions.

Since:

$$\begin{aligned} \eta(0,m) &= \xi(0,m) - 0 = 0 \\ \eta(m,m) &= \xi(m,m) - \mu(m)m = -\ln P(0,m) + \ln P(0,m) = 0 \end{aligned}$$

the stochastic part of the bond return is constrained at $t = 0$ and $t = m$.

6.2. Brownian bridge process

Let

$$\eta(t,m) = \sigma\varepsilon(t/m) \tag{6.3}$$

where $\{\varepsilon(s) : s \in [0,1]\}$ is a Brownian Bridge process and σ is the instantaneous standard deviation of the excess return.

$\{\varepsilon(s) : s \in [0,1]\}$ is a standardised Brownian Bridge process if:

- $P(\varepsilon(0) = 0) = 1$,
- the process is Gaussian, where

$$\begin{aligned} \mathbb{E}[\varepsilon(s)] &= 0 & \forall\, s \in [0,1] \\ \text{and } \mathbb{E}[\varepsilon(s_i)\varepsilon(s_j)] &= s_i(1 - s_j) & 0 \leq s_i \leq s_j \leq 1 \end{aligned}$$

- $P(\varepsilon(1) = 0) = 1$.

Hence the standardised Brownian Bridge process is an augmented standardised Brownian motion[1] with the added requirement that it takes on the value zero at time $s = 1$. The definition of $\eta(t,m)$ in (6.3) satisfies the requirements

[1] The stochastic process $\{Z(s) : s \in [0,\infty)\}$ is a standardised Brownian motion if:

- $P(Z(0) = 0) = 1$,
- the process is Gaussian with

$$\begin{aligned} \mathbb{E}[Z(s)] &= 0 & \forall\, s \geq 0 \\ \text{and } \mathbb{E}[Z(s_i)Z(s_j)] &= s_i & 0 \leq s_i \leq s_j \end{aligned}$$

that $\eta(0,m) = \eta(m,m) = 0$ and hence it may be specified as a Brownian Bridge process.

Our assumption that markets are efficient implies that all currently known information is included in current market prices, or market yields. Under this assumption, the return on the bond can only vary from the deterministic return $\mu(m)t$, by unanticipated information becoming known. Had this unanticipated information been known, market yields would have adjusted to accommodate it. Hence we may conclude that:

$$\mathbb{E}\left[\eta(t,m)\right] = 0$$

It is appropriate to represent the unexpected returns by means of a Gaussian process, since they are a result of random economic events.

Let $\{Z(s) : s \in [0,\infty)\}$ be a standardised Brownian motion. The properties of a standardised Brownian motion process, for $t \geq 0$ and $\Delta \geq 0$ where $0 \leq t + \Delta < 1$, imply[2]:

$$\mathbb{E}\left[Z\left(\frac{t+\Delta}{1-t-\Delta}\right) \middle| Z\left(\frac{t}{1-t}\right) = z \right] = z$$

$$\mathbb{E}\left[\left\{ Z\left(\frac{t+\Delta}{1-t-\Delta}\right) - Z\left(\frac{t}{1-t}\right) \right\}^2 \middle| Z\left(\frac{t}{1-t}\right) = z \right] = \frac{\Delta}{(1-t)(1-t-\Delta)}$$

Define a diffusion process $W(t)$, as follows:

---

[2]Consider a Brownian motion B_t starting at x, that is $P(B_0 = x) = 1$. Then the following are true [**43**]:

$$\begin{aligned}
\mathbb{E}\left[B_t \middle| B_0 = x\right] &= x \qquad \forall\, t \geq 0 \\
\mathbb{E}\left[(B_t - B_0)^2 \middle| B_0 = x\right] &= t \\
\mathbb{E}\left[(B_t - B_0)(B_s - B_0) \middle| B_0 = x\right] &= s \wedge t
\end{aligned}$$

Hence:

$$\begin{aligned}
\mathbb{E}[(B_t - B_s)^2 | B_0 = x] &= \mathbb{E}[(B_t - B_0)^2 - 2(B_t - B_0)(B_s - B_0) \\
&\quad + (B_s - B_0)^2 | B_0 = x] \quad s \leq t \\
&= t - 2s + s \\
&= t - s
\end{aligned}$$

and translating this analysis into the required notation, we may calculate

$$\begin{aligned}
\mathbb{E}\left[\left\{ Z\left(\frac{t+\Delta}{1-t-\Delta}\right) - Z\left(\frac{t}{1-t}\right) \right\}^2 \middle| Z\left(\frac{t}{1-t}\right) = z \right] &= \frac{t+\Delta}{1-t-\Delta} - \frac{t}{1-t} \\
&= \frac{\Delta}{(1-t)(1-t-\Delta)}
\end{aligned}$$

$$W(t) = (1-t) Z\Big(\frac{t}{1-t}\Big) \qquad 0 \le t < 1 \tag{6.4}$$

Hence:

- $P(W(0) = 0) = 1$,
- $P(W(1) = 0) = 1$,
- and

$$\begin{aligned}
&\mathbb{E}\left[W(t)\right] = \ 0 && \forall\, t \in [0,1] \\
&\mathbb{E}\left[W(t_i)W(t_j)\right] \\
&= (1-t_i)(1-t_j)\,\mathbb{E}\left[Z\Big(\frac{t_i}{1-t_i}\Big)Z\Big(\frac{t_j}{1-t_j}\Big)\right] \\
&= (1-t_i)(1-t_j)\Big(\frac{t_i}{1-t_i}\Big) \\
&= t_i(1-t_j) && 0 \le t_i \le t_j \le 1
\end{aligned}$$

which implies that $\{W(t) : t \in [0,1]\}$ is a standardised Brownian Bridge process.

If $\{X(t) : t \in T\}$ is a diffusion process, where T is some index set, it may be represented as

$$dX = \mu_X(x,t)dt + \sigma_X(x,t)dZ \qquad t \in T$$

where $\mu_X(x,t)$ and $\sigma_X(x,t)$ are respectively, the instantaneous mean and variance of the diffusion process. More specifically, we have:

$$\mu_X(x,t) = \lim_{\Delta \to 0} \frac{\mathbb{E}\left[\,X(t+\Delta) - X(t)|\, X(t) = x\right]}{\Delta} \qquad \text{and} \tag{6.5a}$$

$$\sigma^2_X(x,t) = \lim_{\Delta \to 0} \frac{\mathbb{E}\left[\left\{X(t+\Delta) - X(t)\right\}^2 \Big| \, X(t) = x\right]}{\Delta} \tag{6.5b}$$

Hence to determine the process describing the evolution of the Brownian Bridge we need to calculate the instantaneous mean and variance, $\mu_W(w,t)$ and $\sigma^2_W(w,t)$ respectively. Using the specifications in (6.5) above, we have:

$$\begin{aligned}
&\mu_W(w,t) \\
&= \lim_{\Delta \to 0} \frac{\mathbb{E}\left[W(t+\Delta) - W(t)\big|W(t) = w\right]}{\Delta} \\
&= \lim_{\Delta \to 0} \frac{\mathbb{E}\left[(1-t-\Delta)Z\Big(\frac{t+\Delta}{1-t-\Delta}\Big) - (1-t)Z\Big(\frac{t}{1-t}\Big)\Big|\, Z\Big(\frac{t}{1-t}\Big) = \frac{w}{1-t}\right]}{\Delta} \\
&= \lim_{\Delta \to 0} \frac{(1-t-\Delta)\frac{w}{1-t} - w}{\Delta} \\
&= -\frac{w}{1-t}
\end{aligned}$$

and

$$\begin{aligned}
&\sigma_W^2(w,t) \\
&= \lim_{\Delta\to 0} \frac{\mathbb{E}\left[\left\{W(t+\Delta)-W(t)\right\}^2 \middle| W(t)=w\right]}{\Delta} \\
&= \lim_{\Delta\to 0} \frac{(1-t-\Delta)^2 \ \mathbb{E}\left[\left\{Z\left(\frac{t+\Delta}{1-t-\Delta}\right)-\frac{1-t}{1-t-\Delta}\,Z\left(\frac{t}{1-t}\right)\right\}^2 \middle| Z\left(\frac{t}{1-t}\right)=\frac{w}{1-t}\right]}{\Delta}
\end{aligned}$$

Now:

$$\begin{aligned}
&\mathbb{E}\left[Z^2\left(\frac{t+\Delta}{1-t-\Delta}\right)\middle|\ Z\left(\frac{t}{1-t}\right)=\frac{w}{1-t}\right] \\
&=\mathbb{E}\Bigg[\left(\ Z\left(\frac{t+\Delta}{1-t-\Delta}\right)-Z\left(\frac{t}{1-t}\right)\right)^2 \\
&\qquad +2\,Z\left(\frac{t}{1-t}\right)Z\left(\frac{t+\Delta}{1-t-\Delta}\right)-Z^2\left(\frac{t}{1-t}\right)\Bigg|\ Z\left(\frac{t}{1-t}\right)=\frac{w}{1-t}\Bigg] \\
&=\frac{\Delta}{(1-t)(1-t-\Delta)}+\frac{w^2}{(1-t)^2}
\end{aligned}$$

and so:

$$\begin{aligned}
&\mathbb{E}\left[\left\{\ Z\left(\frac{t+\Delta}{1-t-\Delta}\right)-\frac{1-t}{1-t-\Delta}\,Z\left(\frac{t}{1-t}\right)\right\}^2\middle|\ Z\left(\frac{t}{1-t}\right)=\frac{w}{1-t}\right] \\
&= \ \mathbb{E}\left[\ Z^2\left(\frac{t+\Delta}{1-t-\Delta}\right)\middle|\ Z\left(\frac{t}{1-t}\right)=\frac{w}{1-t}\right] \\
&\quad -\frac{2w}{1-t-\Delta}\ \mathbb{E}\left[\ Z\left(\frac{t+\Delta}{1-t-\Delta}\right)\middle|\ Z\left(\frac{t}{1-t}\right)=\frac{w}{1-t}\right]+\frac{w^2}{(1-t-\Delta)^2} \\
&=\frac{\Delta}{(1-t)(1-t-\Delta)}+\frac{w^2}{(1-t)^2}-\frac{2w^2}{(1-t)(1-t-\Delta)}+\frac{w^2}{(1-t-\Delta)^2} \\
&=\frac{\Delta}{(1-t)(1-t-\Delta)}+\frac{\Delta^2w^2}{(1-t)^2(1-t-\Delta)^2}
\end{aligned}$$

Hence:

$$\begin{aligned}
\sigma_W^2(w,t) &= \lim_{\Delta\to 0}\frac{(1-t-\Delta)^2\left\{\frac{\Delta}{(1-t)(1-t-\Delta)}+\frac{\Delta^2w^2}{(1-t)^2(1-t-\Delta)^2}\right\}}{\Delta} \\
&= \lim_{\Delta\to 0}\frac{1-t-\Delta}{1-t}+\frac{\Delta w^2}{(1-t)^2} \\
&= 1
\end{aligned}$$

We have determined the instantaneous mean and variance of the standardised Brownian Bridge process to be $\mu_W(w,t) = -\frac{w}{1-t}$ and $\sigma_W^2(w,t) = 1$ respectively. Hence the standardised Brownian Bridge process is characterised as:

$$dW = -\frac{w}{1-t}\,dt + dZ \qquad 0 \leq t < 1 \tag{6.6}$$

The standardised Brownian Bridge is subject to a restoring force, pulling it back towards zero. The instantaneous variance remains constant, but the total variance, as at time t, is non-stationary and expressed as $\mathbb{E}\left[W^2(t)\right] = t(1-t)$. This non-stationarity is due to the imposed terminal constraint. Since the excess return is assumed to have the functional form of (6.3), it is modelled as:

$$d\eta(t,m) = \frac{-\eta(t,m)}{1-t}\,dt + \sigma dZ \qquad 0 \leq t < 1$$

The economic interpretation of the time t return $\xi(t,m)$, is given by (6.1) and (6.2); hence :

$$\begin{aligned}
\exp\xi(t,m) &= \frac{P(t,m)}{P(0,m)} \\
\Rightarrow P(t,m) &= P(0,m)\exp\left(\mu(m)t + \eta(t,m)\right)
\end{aligned}$$

The discount bond price dynamics are determined by application of Ito's Lemma:

$$\begin{aligned}
dP &= \frac{\partial P}{\partial \eta}\,d\eta + \frac{1}{2}\frac{\partial^2 P}{\partial \eta^2}\,d\eta d\eta + \frac{\partial P}{\partial t}\,dt \\
&= P(t,m)\left(-\frac{\eta(t,m)}{1-t}\,dt + \sigma dZ\right) + \tfrac{1}{2}P(t,m)\sigma^2 dt + \mu(m)P(t,m)dt \\
\Rightarrow \frac{dP}{P} &= \left(\mu(m) + \tfrac{1}{2}\sigma^2 + \frac{\mu(m)t - \xi(t,m)}{1-t}\right)dt + \sigma dZ \\
&= \left(\mu(m) + \tfrac{1}{2}\sigma^2 + \frac{\mu(m)t + \ln\left[P(0,m)/P(t,m)\right]}{1-t}\right)dt + \sigma dZ \qquad (6.7)
\end{aligned}$$

6.3. Option valuation

We now examine the equilibrium pricing of a European call option where the underlying is the risk-free zero coupon bond described by (6.7). Without loss of generality, the following assumptions can be made:

- bond matures as time 1,
- option expiry is at time τ where $\tau \leq 1$,

- option exercise price is $k \leq 1$.

Further we assume:

- Markets are frictionless. That is:
  - no transaction costs, no taxes,
  - continuous trading,
  - unlimited lending and borrowing,
  - unlimited short sales.
- The dynamics of the underlying risk-free zero coupon bond are described by (6.7), which is a non-standard Brownian Bridge process.
- There exists a risk-free discount bond maturing at time τ, also described by (6.7). Assume that:

  $$\mathbb{E}\left[dZ(t,1)\; dZ(t,\tau)\right] = \rho\; dt$$

  where ρ is the instantaneous correlation between the unexpected returns on bonds maturing at time 1 and τ respectively. An initial upward sloping yield curve is assumed, that is $P(0,\tau) > P(0,1)$, which implies that initial forward rates are positive.
- Investors are assumed to be rational and to have consensus views on the instantaneous standard deviations of bond returns and their distributions. They need not have the same views of the term structure or expected returns on bonds of various maturities.

By forming a hedge portfolio in bonds and the call, the equilibrium value of a European call may be calculated. The time t price of a European call on a discount bond with maturity 1, expiry τ and strike k is:

$$C(k,\tau) = P(t,1)N(h_1) - P(t,\tau)kN(h_2) \tag{6.8}$$

where

$$\begin{aligned} h_1 &= \frac{\ln\left(P(t,1)/k\right) - \ln P(t,\tau) + \left(\nu_\tau^2/2\right)(\tau - t)}{\nu_\tau\sqrt{\tau - t}} \\ h_2 &= h_1 - \nu_\tau\sqrt{\tau - t} \\ \nu_\tau^2 &= \sigma_1^2 + \sigma_\tau^2 - 2\rho\sigma_1\sigma_\tau \end{aligned}$$

and $N(\cdot)$ is the cumulative normal distribution. Here ν_τ is the volatility, at time τ of the price of the bond with maturity time 1, i.e. it is the forward bond volatility. Consider the forward bond price $P(\tau,1) = P(t,1)/P(t,\tau)$. Its volatility may be calculated as:

$$\nu_\tau = \sqrt{\sigma_1^2 + \sigma_\tau^2 - 2\rho\sigma_1\sigma_\tau}$$

where σ_1 and σ_τ are the current (time t) volatilities of the bonds with maturity time 1 and τ respectively.

6.4. Conclusion

In this model, BT assume that default-free discount bond prices follow a non-standardised Brownian Bridge process. This assumption satisfies the pull-to-par characteristic of bond prices, but ignores the dependence of bond prices on the underlying interest rate term structure. No restrictions are placed on the dynamics of the bond price to ensure a model that precludes profitable arbitrage. This essential factor limits the usefulness and applicability of this model.

CHAPTER 7

The Hull and White Model

The Vasicek [**50**] and CIR [**18**] models, studied in Chapters 1 and 2 respectively, allow all interest rate contingent claims to be valued in a consistent manner, but involve unobservable parameters and do not provide a perfect fit for the current interest rate term structure.

However, the process describing the evolution of the short-term interest rate may be deduced from the observed term structure of interest rates and interest rate volatilities. Hence the Vasicek and CIR models may be extended so as to be consistent with the current term structure of interest rates and the current spot interest rate volatilities or current forward rate volatilities.

7.1. General model formulation

The Vasicek and CIR models are special cases of a general mean reverting process of the form:

$$dr = a(b-r)dt + \sigma r^{\beta} dz \tag{7.1}$$

where $\beta = 0$ for the Vasicek model and $\beta = \frac{1}{2}$ for the CIR model.

Since market expectations of interest rate movements can be time-dependent, the drift and volatility parameters should be functions of time:

$$dr = \left[\theta(t) + a(t)(b-r)\right] dt + \sigma(t) r^{\beta} dz \tag{7.2}$$

where $\theta(t)$ is the drift rate imposed on the interest rate which otherwise reverts to a constant level b. Rewriting (7.2) in the form:

$$dr = a(t)\left[\left(\frac{\theta(t)}{a(t)} + b\right) - r\right] dt + \sigma(t) r^{\beta} dz \tag{7.3}$$

gives a mean reverting model where the reversion level is a function of time.

Hull and White (HW) [**28**] make assumptions about the market price of interest rate risk and fit the Vasicek and CIR special cases of the above mean reverting model to the current term structure of interest rates and spot (or forward) interest rate volatilities.

7.2. Extension of the Vasicek model

HW propose an extension to the Vasicek model of the form:

$$dr = [\theta(t) + a(t)(b - r)]\, dt + \sigma(t)\, dz \tag{7.4}$$

This is (7.2) with $\beta = 0$. Assuming that the market price of interest rate risk has the functional form $\lambda(t)$, and is bounded on any time interval $(0, \tau)$, we may apply Ito's Lemma to derive the general partial differential equation that must be satisfied by any interest rate contingent claim, g:

$$\begin{aligned} \frac{dg}{g} &= \left(\frac{g_t}{g} + \left(\frac{\theta(t) + a(t)(b - r)}{g} \right) g_r + \frac{\sigma^2(t) g_{rr}}{2g} \right) dt + \frac{\sigma(t) g_r}{g}\, dz \\ &= \mu(t)dt + s(t)dz \end{aligned}$$

The market price of risk represents the excess return required above the risk-free rate. This relationship is denoted by:

$$\mu(t) - r = \lambda(t)s(t)$$

and so

$$\begin{aligned} g_t + (\theta(t) + a(t)(b - r))\, g_r + \tfrac{1}{2}\, \sigma^2(t) g_{rr} - rg &= \lambda(t)\sigma(t)g \\ \Rightarrow g_t + (\phi(t) - a(t)r)\, g_r + \tfrac{1}{2}\, \sigma^2(t) g_{rr} - rg &= 0 \end{aligned} \tag{7.5}$$

where $\phi(t) = a(t)b + \theta(t) - \lambda(t)\sigma(t)$. Assume that the price of the contingent claim g, has the form:

$$g(r, t, T) = A(t, T) e^{-B(t,T)r}, \tag{7.6}$$

with boundary condition $g(r, T, T) = 1$. Now:

$$\begin{aligned} g_t &= A_t\, e^{-Br} - AB_t\, r\, e^{-Br} \\ g_r &= -AB\, e^{-Br} \\ g_{rr} &= AB^2\, e^{-Br} \end{aligned}$$

Substituting into (7.5):

$$\begin{aligned} A_t\, e^{-Br} - AB_t\, r\, e^{-Br} - AB\, (\, \phi(t) - a(t)r\,)\, e^{-Br} & \\ + \tfrac{1}{2}\, AB^2\, \sigma^2(t)\, e^{-Br} - A\, r\, e^{-Br} &= 0 \end{aligned}$$

$$e^{-Br} \left(A_t - AB\phi(t) + \tfrac{1}{2}\, AB^2 \sigma^2(t) \right) + A\, r\, e^{-Br}\, (Ba(t) - B_t - 1) = 0 \tag{7.7}$$

Therefore to solve (7.7) we must solve the system of simultaneous equations:

$$
\begin{aligned}
&(7.8a) \qquad A_t - AB\phi(t) + \tfrac{1}{2}AB^2\sigma^2(t) = 0 \qquad \text{with } A(T,T) = 1 \\
&(7.8b) \qquad Ba(t) - B_t - 1 = 0 \qquad \text{with } B(T,T) = 0
\end{aligned}
$$

For $\phi(t)$, $\sigma(t)$ and $a(t)$ constant, (7.8a) and (7.8b) are solved to yield the Vasicek model where the bond price has the form assumed in (7.6) with[1]:

$$
\begin{aligned}
&(7.9) \quad A(t,T) = \exp\left[\frac{(B(t,T) - T + t)\left(a\,\phi - \sigma^2/2\right)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}\right] \\
&(7.10) \quad B(t,T) = \frac{1}{a}\left(1 - e^{-a(T-t)}\right)
\end{aligned}
$$

For the extended, time-dependent model, $\sigma(t)$ should be chosen to reflect the current and future volatilities of the short-term interest rate. $A(0,T)$ and $B(0,T)$ are coefficients associated with the current term structure and are hence functions of the current interest rate term structure, current term structure of spot/forward interest rate volatilities, and $\sigma(0)$ (the current volatility of the short-term interest rate). Since the current term structure is observable, we are able to determine $A(0,T)$, $B(0,T)$ and $\sigma(t)$. Therefore, we must determine $A(t,T)$, $B(t,T)$, $a(t)$ and $\phi(t)$ in terms of $A(0,T)$, $B(0,T)$ and $\sigma(t)$.

First, differentiate (7.8a) and (7.8b) with respect to T. From (7.8a) we have:

$$A_{tT} - \phi(t)A_T B - \phi(t)AB_T + \tfrac{1}{2}\sigma^2(t)A_T B^2 + \sigma^2(t)ABB_T = 0$$

Also, from (7.8a):

$$\phi(t) = \tfrac{1}{2}\sigma^2(t)B + \frac{A_t}{AB} \tag{7.11}$$

Hence:

$$
\begin{aligned}
A_{tT} - \left[\tfrac{1}{2}\sigma^2(t)B + \frac{A_t}{AB}\right](A_T B + AB_T) + \tfrac{1}{2}\sigma^2(t)A_T B^2 + \sigma^2(t)ABB_T &= 0 \\
A_{tT} - \tfrac{1}{2}\sigma^2(t)A_T B^2 - \tfrac{1}{2}\sigma^2(t)ABB_T - \frac{A_t A_T}{A} - \frac{A_t B_T}{B} & \\
+\tfrac{1}{2}\sigma^2(t)A_T B^2 + \sigma^2(t)ABB_T &= 0 \\
(7.12) \qquad \Rightarrow \qquad ABA_{tT} - BA_t A_T - AA_t B_T + \tfrac{1}{2}\sigma^2(t)A^2B^2B_T &= 0
\end{aligned}
$$

[1]These are the same formulae as calculated for the Vasicek model in Chapter 1 equations (1.25) and (1.26) with the following notational substitutions:

$$\lambda = -q, \quad \sigma = \rho, \quad a = \alpha, \quad \phi/a = \gamma + \rho q/\alpha.$$

with $A(T,T) = 1$ and $A(0,T) = \xi$ where ξ is some known value. Similarly differentiating (7.8b) with respect to T yields:

$$B_{tT} - a(t)B_T = 0$$

From (7.8b) we have:

$$a(t) = \frac{B_t}{B} + \frac{1}{B} \tag{7.13}$$

Therefore:

$$\begin{aligned} B_{tT} - \frac{B_t B_T}{B} - \frac{B_T}{B} &= 0 \\ \Rightarrow \qquad BB_{tT} - B_t B_T - B_T &= 0 \end{aligned} \tag{7.14}$$

with $B(T,T) = 0$ and $B(0,T) = \eta$ where η is some known value. HW [**28**] solve (7.12) and (7.14) to yield:

$$B(t,T) = \frac{B(0,T) - B(0,t)}{\frac{\partial B(0,t)}{\partial t}} \tag{7.15}$$

$$\begin{aligned} \hat{A}(t,T) &= \hat{A}(0,T) - \hat{A}(0,t) - B(t,T)\frac{\partial \hat{A}(0,t)}{\partial t} \\ &\quad -\frac{1}{2}\left[B(t,T)\frac{\partial B(0,t)}{\partial t}\right]^2 \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \end{aligned} \tag{7.16}$$

$$\text{where} \qquad \hat{A}(t,T) = \log A(t,T)$$

We have solved for $A(t,T)$ and $B(t,T)$ in terms of the initial term structure. Now solve for $a(t)$ and $\phi(t)$. Differentiating (7.15) yields:

$$\begin{aligned} B_t &= \frac{\left(-\left(\frac{\partial B(0,t)}{\partial t}\right)^2 - (B(0,T) - B(0,t))\frac{\partial^2 B(0,t)}{\partial t^2}\right)}{\left(\frac{\partial B(0,t)}{\partial t}\right)^2} \\ \Rightarrow \frac{B_t}{B} &= \frac{\left(-\left(\frac{\partial B(0,t)}{\partial t}\right)^2 - (B(0,T) - B(0,t))\frac{\partial^2 B(0,t)}{\partial t^2}\right)}{\left(\frac{\partial B(0,t)}{\partial t}(B(0,T) - B(0,t))\right)} \end{aligned}$$

and so from (7.13) we solve for $a(t)$ as:

$$
\begin{aligned}
a(t) &= -\frac{\frac{\partial B(0,t)}{\partial t}}{B(0,T)-B(0,t)} - \frac{\left(\frac{\partial^2 B(0,t)}{\partial t^2}\right)}{\left(\frac{\partial B(0,t)}{\partial t}\right)} + \frac{\frac{\partial B(0,t)}{\partial t}}{B(0,T)-B(0,t)} \\
&= -\left(\frac{\partial^2 B(0,t)}{\partial t^2}\right) \Big/ \left(\frac{\partial B(0,t)}{\partial t}\right) \qquad (7.17)
\end{aligned}
$$

Differentiating (7.16) yields:

$$
\begin{aligned}
\hat{A}_t = &-\frac{\partial \hat{A}(0,t)}{\partial t} - B_t \frac{\partial \hat{A}(0,t)}{\partial t} - B\frac{\partial^2 \hat{A}(0,t)}{\partial t^2} \\
&- B\frac{\partial B(0,t)}{\partial t}\left[B_t \frac{\partial B(0,t)}{\partial t} + B\frac{\partial^2 B(0,t)}{\partial t^2}\right] \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&- \frac{1}{2}\left[B\frac{\partial B(0,t)}{\partial t}\right]^2 \left[\frac{\sigma(t)}{\frac{\partial B(0,t)}{\partial t}}\right]^2
\end{aligned}
$$

Since $\hat{A}(t,T) = \log A(t,T)$ we have:

$$A_t = \hat{A}_t \exp \hat{A}(t,T) = \hat{A}_t A$$

Hence

$$
\begin{aligned}
\frac{A_t}{AB} &= \frac{\hat{A}_t}{B} = -\frac{\frac{\partial \hat{A}(0,t)}{\partial t}}{B} - \frac{B_t}{B}\frac{\partial \hat{A}(0,t)}{\partial t} - \frac{\partial^2 \hat{A}(0,t)}{\partial t^2} \\
&\quad -\frac{\partial B(0,t)}{\partial t}\left[B_t \frac{\partial B(0,t)}{\partial t} + B\frac{\partial^2 B(0,t)}{\partial t^2}\right] \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&\quad -\frac{1}{2}B\sigma^2(t)
\end{aligned}
$$

Therefore substituting into (7.11) we solve for $\phi(t)$ in terms of the initial term structure:

$$
\begin{aligned}
\phi(t) &= -\frac{\frac{\partial \hat{A}(0,t)}{\partial t}}{B} - \frac{B_t}{B}\frac{\partial \hat{A}(0,t)}{\partial t} - \frac{\partial^2 \hat{A}(0,t)}{\partial t^2} \\
&\quad -B_t\left(\frac{\partial B(0,t)}{\partial t}\right)^2 \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&\quad -B\frac{\partial B(0,t)}{\partial t}\frac{\partial^2 B(0,t)}{\partial t^2} \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau
\end{aligned}
$$

Also, from (7.8b) and (7.17) we have

$$\begin{aligned}
\frac{B_t}{B} &= a(t) - \frac{1}{B} \\
\text{and} \quad B_t\left(\frac{\partial B(0,t)}{\partial t}\right)^2 &= -B\frac{\partial B(0,t)}{\partial t}\frac{\partial^2 B(0,t)}{\partial t^2} - \left(\frac{\partial B(0,t)}{\partial t}\right)^2
\end{aligned}$$

Hence

$$\begin{aligned}
\phi(t) &= -\frac{\frac{\partial \hat{A}(0,t)}{\partial t}}{B} - a(t)\frac{\partial \hat{A}(0,t)}{\partial t} + \frac{\frac{\partial \hat{A}(0,t)}{\partial t}}{B} - \frac{\partial^2 \hat{A}(0,t)}{\partial t^2} \\
&\quad + B\frac{\partial B(0,t)}{\partial t}\frac{\partial^2 B(0,t)}{\partial t^2}\int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&\quad + \left(\frac{\partial B(0,t)}{\partial t}\right)^2 \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&\quad - B\frac{\partial B(0,t)}{\partial t}\frac{\partial^2 B(0,t)}{\partial t^2}\int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&= -a(t)\frac{\partial \hat{A}(0,t)}{\partial t} - \frac{\partial^2 \hat{A}(0,t)}{\partial t^2} + \left(\frac{\partial B(0,t)}{\partial t}\right)^2 \int_0^t \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau
\end{aligned}$$

and we have specified all the required model parameters in terms of the initial yield curve.

7.3. Pricing contingent claims within the extended Vasicek framework

Let $P(r,t,T)$ be the time t price of a discount bond maturing at time T. Since this is an interest rate contingent claim, it may be written in the functional form specified in (7.6):

$$P(r,t,T) = A(t,T)e^{-B(t,T)r} \tag{7.18}$$

By Ito's Lemma we have:

$$\begin{aligned}
dP &= \frac{\partial P}{\partial t}\,dt + \frac{\partial P}{\partial r}\,dr + \tfrac{1}{2}\,\frac{\partial^2 P}{\partial r^2}\,drdr \\
&= \frac{\partial P}{\partial t}\,dt - ABe^{-Br}\left[(\theta(t) + a(t)(b-r))\,dt + \sigma(t)dz\right] \\
&\quad + \tfrac{1}{2}\,AB^2e^{-Br}\sigma^2(t)dt \\
&= P_t dt - BP\left[\theta(t) + a(t)(b-r)\right]dt - BP\sigma(t)dz + \tfrac{1}{2}\,B^2P\sigma^2(t)dt
\end{aligned}$$

Hence the price process of the discount bond is described by the stochastic equation:

$$dP = \left[P_t - BP\left(\theta(t) + a(t)(b-r)\right) + \tfrac{1}{2}\,B^2 P \sigma^2(t)\right] dt - PB\sigma(t)dz \tag{7.19}$$

The relative volatility of $P(r,t,T)$ is $B(t,T)\sigma(t)$. Since it is independent of r, the distribution of the bond price at any time t^*, conditional on its value at an earlier time $\hat{t}$, must be lognormally distributed.

Consider a European option on this discount bond. This option has the following characteristics:

X – exercise price,
T – option expiry time,
s – bond maturity time,
t – current (valuation) time, where $t \le T \le s$.

This option may be viewed as being equivalent to an option to exchange X units of a discount bond maturing at time T for one unit of a discount bond maturing at time s. Define:

$\alpha_1(\tau)$ – time τ volatility of the price of a discount bond with maturity s,
$\alpha_2(\tau)$ – time τ volatility of the price of a discount bond with maturity T,
$\rho(\tau)$ – instantaneous correlation between the bond prices,

Then the price of a European call option may be written as[2]:

$$C = P(r,t,s)N(h) - XP(r,t,T)N(h - \sigma_{P_f(T,s)}) \tag{7.20}$$

$$\text{where} \quad h = \frac{\ln\left(P(r,t,s)/\left(XP(r,t,T)\right)\right)}{\sigma_{P_f(T,s)}} + \tfrac{1}{2}\,\sigma_{P_f(T,s)} \tag{7.21}$$

$$\text{and } \sigma^2_{P_f(T,s)} = \int_t^T \left[\alpha_1^2(\tau) + \alpha_2^2(\tau) - 2\rho(\tau)\alpha_1(\tau)\alpha_2(\tau)\right] d\tau \tag{7.22}$$

One of the characteristics of a one factor model is that instantaneous returns on bonds of all maturities are perfectly correlated. Hence $\rho(\tau) = 1$ for all τ. Also, from the equation of the general bond price process (7.19) we may write the volatilities of the two bonds as:

[2]Here, the appropriate volatility to use is that of the forward bond price, i.e. the volatility of the time T price of the bond maturing at time s which may be expressed as $P_f(T,s) = \frac{P(r,t,s)}{P(r,t,T)}$. Use Ito's Lemma to determine this volatility:

$$\begin{aligned} \alpha_1(\tau) &= \sigma(\tau)B(\tau,s) \\ \alpha_2(\tau) &= \sigma(\tau)B(\tau,T) \end{aligned} \tag{7.23}$$

7.3.1. Time-dependent parameters. Given the functional form of the volatilities above, calculate the volatility required for option valuation as:

$$\begin{aligned} \sigma^2_{P_f(T,s)} &= \int_t^T \left[\sigma^2(\tau)B^2(\tau,s) + \sigma^2(\tau)B^2(\tau,T) - 2\sigma^2(\tau)B(\tau,s)B(\tau,T)\right] d\tau \\ &= \int_t^T \sigma^2(\tau)\left[B(\tau,s) - B(\tau,T)\right]^2 d\tau \end{aligned}$$

From equation (7.15) we have $B(t,T) = (B(0,T) - B(0,t)) / \left(\frac{\partial B(0,t)}{\partial t}\right)$ and so:

$$\begin{aligned} \sigma^2_{P_f(T,s)} &= \int_t^T \sigma^2(\tau)\left[\frac{B(0,s) - B(0,\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}} - \frac{B(0,T) - B(0,\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\ &= \left[B(0,s) - B(0,T)\right]^2 \int_t^T \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \end{aligned} \tag{7.24}$$

$$\begin{aligned} dP_f(T,s) &= \frac{\partial P_f(T,s)}{\partial t}dt + \frac{\partial P_f(T,s)}{\partial P(r,t,s)}dP(r,t,s) + \frac{\partial P_f(T,s)}{\partial P(r,t,T)}dP(r,t,T) \\ &\quad + \frac{1}{2}\frac{\partial^2 P_f(T,s)}{\partial P(r,t,s)^2}dP(r,t,s)dP(r,t,s) + \frac{1}{2}\frac{\partial^2 P_f(T,s)}{\partial P(r,t,T)^2}dP(r,t,T)dP(r,t,T) \\ &\quad + \frac{\partial^2 P_f(T,s)}{\partial P(r,t,s)\partial P(r,t,T)}dP(r,t,s)dP(r,t,T) \\ &= \frac{\partial P_f(T,s)}{\partial t}dt + \frac{1}{P(r,t,T)}\left(\mu_{P(r,t,s)}dt - P(r,t,s)\alpha_1(t)dz\right) \\ &\quad - \frac{P(r,t,s)}{P^2(r,t,T)}\left(\mu_{P(r,t,T)}dt - P(r,t,T)\alpha_2(t)dz\right) \\ &\quad + \frac{P(r,t,s)}{P^3(r,t,T)}P^2(r,t,T)\alpha_2^2(t)dt - \frac{1}{P^2(r,t,T)}P(r,t,T)P(r,t,s)\rho\,\alpha_1(t)\alpha_2(t)dt \\ &= \mu_{P_f(T,s)}dt - P_f(T,s)\left(\alpha_1(t) - \alpha_2(t)\right)dz \end{aligned}$$

Hence, the instantaneous volatility of the forward bond price is $\alpha_1(t) - \alpha_2(t)$ and so the square of forward price volatility over the life of the option is:

$$\begin{aligned} \sigma^2_{P_f(T,s)} &= \int_t^T \left[\alpha_1(\tau) - \alpha_2(\tau)\right]^2 d\tau \\ &= \int_t^T \left[\alpha_1^2(\tau) + \alpha_2^2(\tau) - 2\rho(\tau)\alpha_1(\tau)\alpha_2(\tau)\right] d\tau \end{aligned}$$

Equations (7.20), (7.21) and (7.24) give analytical formulae for the price of a European call option on a discount bond. The corresponding European put option price may be obtained via put-call parity.

This formulation of the pricing formula is very attractive since $a(t)$ and $\sigma(t)$ may be chosen in such a way that a whole set of cap or swaption prices observed at time 0 can be exactly replicated. However, this full level of precision results in undesirable side effects [**45**], [**29**]. Examine the process for the short-term interest rate as considered thus far:

$$dr = [\theta(t) + a(t)(b - r)]\,dt + \sigma(t)dz$$

Here $\theta(t)$ is chosen such that the prices of all discount bonds at the initial time are reproduced, i.e. the initial term structure is matched, $a(t)$ and $\sigma(t)$ provide another two degrees of freedom which allow matching of the initial volatility term structure and volatilities of the short-term interest rate in the future. Initial volatilities of rates depend on $\sigma(0)$ and $a(t)$, hence $a(t)$ defines the relative volatilities of long- and short-term interest rates. Finally, $\sigma(t)$ determines the future volatilities of the short-term interest rate.

It is appealing to take advantage of all these degrees of freedom since all the available initial market data will be incorporated into the model. Unfortunately this results in non-stationarity of the volatility term structure which could result in the mispricing of instruments contingent on the future, rather than simply the current, volatility structure.

Fitting all the model parameters to initial option prices results in a model which exactly reflects the initial term structure, but also introduces assumptions about the future evolution of the volatility structure. Making use of all the degrees of freedom produces an over-parameterisation of the model. Hull and White [**29**] recommend keeping the parameters a and σ constant. Within this simplified model, observed cap and swaption prices will only be approximated, but the model evolution can be more directly controlled. A stationary volatility term structure is achieved, resulting in robust pricing of more exotic interest rate options.

7.3.2. Constant parameters. If we allow the volatility of the short-term interest rate and the rate of interest rate reversion to be constant, i.e. $\sigma(t) \equiv \sigma$ and $a(t) \equiv a$, we have[3]:

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a} \tag{7.25}$$

[3]This is the value of $B(\tau, \cdot)$ obtained in Chapter 1.

The corresponding function $A(t,T)$ is obtained from (7.16), setting $\sigma(t) \equiv \sigma$ in the integral [**29**]. Making use of (7.15) and (7.25) we have:

$$\begin{aligned}
\ln A(t,T) &= \ln A(0,T) - \ln A(0,t) - B(t,T)\frac{\partial \ln A(0,t)}{\partial t} \\
&\quad -\frac{1}{2}\left[B(t,T)\frac{\partial B(0,t)}{\partial t}\right]^2 \int_0^t \left[\frac{\sigma}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \\
&= \ln \frac{A(0,T)}{A(0,t)} - B(t,T)\frac{\partial \ln A(0,t)}{\partial t} \\
&\quad -\frac{\sigma^2}{2}\left[B(0,T) - B(0,t)\right]^2 \int_0^t \left[\frac{B(\tau,T)}{B(0,T)-B(0,\tau)}\right]^2 d\tau \\
&= \ln \frac{A(0,T)}{A(0,t)} - B(t,T)\frac{\partial \ln A(0,t)}{\partial t} \\
&\quad -\frac{\sigma^2}{2a^2}\left(e^{-at} - e^{-aT}\right)^2 \int_0^t e^{2a\tau}\, d\tau \\
&= \ln \frac{A(0,T)}{A(0,t)} - B(t,T)\frac{\partial \ln A(0,t)}{\partial t} \qquad (7.26) \\
&\quad -\frac{\sigma^2}{4a^3}\left(e^{-at} - e^{-aT}\right)^2 \left(e^{2at} - 1\right)
\end{aligned}$$

Now, to calculate the volatility required for the option pricing, make use of (7.25) to give:

$$\begin{aligned}
B(\tau,s) &= \frac{1 - e^{-a(s-\tau)}}{a} \\
B(\tau,T) &= \frac{1 - e^{-a(T-\tau)}}{a}
\end{aligned}$$

and so:

$$\begin{aligned}
\alpha_1(\tau) &= \sigma B(\tau,s) = \frac{\sigma\left(1 - e^{-a(s-\tau)}\right)}{a} \\
\text{and} \quad \alpha_2(\tau) &= \sigma B(\tau,T) = \frac{\sigma\left(1 - e^{-a(T-\tau)}\right)}{a}
\end{aligned}$$

Substituting into (7.22), find the appropriate volatility $\sigma_{P_f(T,s)}$ as[4]:

$$\begin{aligned}
\sigma^2_{P_f(T,s)} &= \int_t^T \left[\frac{\sigma^2 \left(1 - e^{-a(s-\tau)}\right)^2}{a^2} + \frac{\sigma^2 \left(1 - e^{-a(T-\tau)}\right)^2}{a^2} \right. \\
&\quad \left. - \frac{2\sigma^2 \left(1 - e^{-a(s-\tau)}\right)\left(1 - e^{-a(T-\tau)}\right)}{a^2} \right] d\tau \\
&= \frac{\sigma^2}{a^2} \int_t^T \left[e^{-2a(s-\tau)} + e^{-2a(T-\tau)} - 2e^{-a(s+T-2\tau)} \right] d\tau \\
&= \frac{\sigma^2}{a^2} \left[\frac{1}{2a} e^{-2a(s-\tau)} + \frac{1}{2a} e^{-2a(T-\tau)} - \frac{1}{a} e^{-a(s+T-2\tau)} \right]_{\tau=t}^{\tau=T} \\
&= \frac{\sigma^2}{2a^3} \left(1 + e^{-2a(s-T)} - 2e^{-a(s-T)}\right)\left(1 - e^{-2a(T-t)}\right) \\
&= \frac{\sigma^2}{2a^3} \left(1 - e^{-a(s-T)}\right)^2 \left(1 - e^{-2a(T-t)}\right)
\end{aligned}$$

Hence:

$$\sigma_{P_f(T,s)} = \frac{v(t,T)}{a}\left(1 - e^{-a(s-T)}\right) \tag{7.27}$$

where

$$v(t,T)^2 = \frac{\sigma^2}{2a}\left(1 - e^{-2a(T-t)}\right).$$

7.4. The extended Cox-Ingersoll-Ross model

Setting $\beta = \frac{1}{2}$ in (7.2) leads to the extension of the CIR model proposed by HW:

$$dr = [\theta(t) + a(t)(b - r)]\, dt + \sigma(t)\sqrt{r} dz \tag{7.28}$$

[4] Alternatively we may substitute the appropriate value of $B(0,\cdot)$ into (7.24) to calculate $\sigma_{P_f(T,s)}$ as:

$$\begin{aligned}
\sigma_{P_f(T,s)} &= \frac{1}{a^2}\left((1 - e^{-as}) - (1 - e^{-aT})\right)^2 \int_t^T \left(\frac{\sigma}{e^{-a\tau}}\right)^2 d\tau \\
&= \frac{\sigma^2}{a^2}\left(e^{-aT} - e^{-as}\right)^2 \int_t^T \left(e^{2a\tau}\right) d\tau \\
&= \frac{\sigma^2}{2a^3}\left(e^{-aT} - e^{-as}\right)^2 \left(e^{2aT} - e^{2at}\right) \\
&= \frac{\sigma^2}{2a^3}\left(1 - e^{-a(s-T)}\right)^2 \left(1 - e^{-2a(T-t)}\right)
\end{aligned}$$

The assumption is made that the market price of interest rate risk has the functional form $\lambda(t)\sqrt{r}$ where $\lambda(t)$ is some function bounded on any time interval $(0,\tau)$. Again, let P be the price of a zero coupon bond and make use of Ito's Lemma to derive its price process:

$$\begin{aligned} dP &= \frac{\partial P}{\partial t}\,dt + \frac{\partial P}{\partial r}\,dr + \tfrac{1}{2}\,\frac{\partial^2 P}{\partial r^2}\,drdr \\ &= P_t dt + P_r\left([\theta(t) + a(t)(b-r)]\,dt + \sigma(t)\sqrt{r}dz\right) + \tfrac{1}{2}\,\sigma^2(t)rP_{rr}\,dt \\ &= \mu(t)dt + s(t)dz \end{aligned}$$

where $\mu(\cdot)$ and $s(\cdot)$ are the drift and volatility of P defined as:

$$\begin{aligned} \mu(t) &= P_t + P_r\,[a(t)b + \theta(t) - a(t)r] + \tfrac{1}{2}\,\sigma^2(t)rP_{rr} \\ s(t) &= \sigma(t)\sqrt{r}P_r \end{aligned}$$

Letting $m(t) = \lambda(t)\sqrt{r}$ be the market price of interest rate risk, we have $\mu(t) - rP = m(t)s(t)$ and so:

$$\begin{aligned} P_t + P_r\,[a(t)b + \theta(t) - a(t)r] + \tfrac{1}{2}\,\sigma^2(t)rP_{rr} - rP = \lambda(t)\sqrt{r}\sigma(t)\sqrt{r}P_r \\ \Rightarrow P_t + P_r\,[a(t)b + \theta(t) - a(t)r - \lambda(t)\sigma(t)r] \\ +\tfrac{1}{2}\,\sigma^2(t)rP_{rr} - rP = 0 \end{aligned} \tag{7.29}$$

Let

$$\begin{aligned} \phi(t) &= a(t)b + \theta(t) \\ \psi(t) &= a(t) + \lambda(t)\sigma(t) \end{aligned}$$

and (7.29) may be expressed as:

$$P_t + P_r\,[\phi(t) - \psi(t)r] + \tfrac{1}{2}\,\sigma^2(t)rP_{rr} - rP = 0 \tag{7.30}$$

As in (7.18), the price of the zero coupon bond is assumed to have the functional form:

$$P(r,t,T) = A(t,T)e^{-B(t,T)r}$$

Hence (7.30) becomes:

$$\begin{aligned} A_t e^{-Br} - AB_t re^{-Br} - ABe^{-Br}\,[\phi(t) - \psi(t)r] & \\ +\tfrac{1}{2}\,AB^2 e^{-Br}\sigma^2(t)r - Are^{-Br} &= 0 \\ \Rightarrow \quad A_t - AB\phi(t) + A\left[-B_t + B\psi(t) + \tfrac{1}{2}\,B^2\sigma^2(t) - 1\right] r &= 0 \end{aligned}$$

To solve this partial differential equation we must solve the system of differential equations:

$$A_t - AB\phi(t) = 0 \tag{7.31a}$$

$$\text{and} \quad B_t - B\psi(t) - \tfrac{1}{2}B^2\sigma^2(t) + 1 = 0 \tag{7.31b}$$

subject to the boundary conditions $A(T,T) = 1$ and $B(T,T) = 0$. In the special case where $\phi(t)$, $\psi(t)$ and $\sigma(t)$ are constant, we solve (7.31a) and (7.31b) to give analytical formulae for $A(t,T)$ and $B(t,T)$ as presented in the original CIR paper[5] [**18**]:

$$\begin{aligned} B(t,T) &= \frac{2\left(e^{\gamma(T-t)}-1\right)}{(\gamma+\psi)\left(e^{\gamma(T-t)}-1\right)+2\gamma} \\ A(t,T) &= \left[\frac{2\gamma e^{(\gamma+\psi)(T-t)/2}}{(\gamma+\psi)\left(e^{\gamma(T-t)}-1\right)+2\gamma}\right]^{2\phi/\sigma^2} \end{aligned}$$

where $\gamma = \sqrt{\psi^2 + 2\sigma^2}$. In the HW extension of this model, $\sigma(t)$ needs to be chosen to reflect the observed current and future volatilities of the short-term interest rate. As in the case of the extended Vasicek model, we use:

- $\sigma(0)$, the current i.e. time 0 volatility of the short-term interest rate,
- current interest rate term structure,
- current volatility term structure.

to determine $A(0,T)$ and $B(0,T)$. These initial conditions, together with the already specified boundary conditions allow us to solve equations (7.31a) and (7.31b).

Firstly, differentiating (7.31b) with respect to T, and multiplying by B yields:

$$BB_{tT} - BB_T\psi(t) - B^2B_T\sigma^2(t) = 0 \tag{7.32}$$

Multiplying (7.31b) by B_T and subtracting (7.32) gives:

$$B_tB_T - BB_{tT} + \tfrac{1}{2}B^2B_T\sigma^2(t) + B_T = 0 \tag{7.33}$$

Since this equation cannot be solved analytically, $B(t,T)$ must be obtained via a numerical method such as finite differences. $\sigma(t)$ is a known function obtained from the current and future short-term interest rate volatilities observed in the market. Hence, once $B(t,T)$ is known, equation (7.31b) may be used to determine $\psi(t)$.

[5]These are the same formulae as calculated for the CIR model in Chapter 2 equations (2.36) and (2.38) with the following notational substitutions: $\phi = \kappa\theta$ and $\psi = \kappa + \lambda$.

Next, solve (7.31a) for $A(t,T)$. From (7.31a) we have:

$$\begin{aligned}
\frac{\partial A}{\partial t} &= AB\phi(t) \\
\Rightarrow \frac{dA}{A} = d\ln A(t,T) &= B(t,T)\phi(t)dt \\
\int_0^t d\ln A(s,T) &= \int_0^t B(s,T)\phi(s)ds \\
\ln A(t,T) - \ln A(0,T) = \ln\frac{A(t,T)}{A(0,T)} &= \int_0^t B(s,T)\phi(s)ds \\
\Rightarrow \quad A(t,T) &= A(0,T)\exp\left[\int_0^t B(s,T)\phi(s)ds\right] \qquad (7.34)
\end{aligned}$$

To be able to fully evaluate $A(t,T)$, $\phi(\cdot)$ must be known. Since $A(T,T) = 1$ and $A(0,T)$ is known, evaluate (7.34) at $t = T$:

$$\begin{aligned}
1 &= A(0,T)\exp\left[\int_0^T B(s,T)\phi(s)ds\right] \\
\Rightarrow \quad -\ln A(0,T) &= \int_0^T B(s,T)\phi(s)ds
\end{aligned}$$

which may be used to determine $\phi(\cdot)$ iteratively. This is computationally time-consuming since $\phi(s)$ needs to be evaluated at each time point in the interval $[0,t]$ so that the integral in (7.34) may be calculated numerically.

7.5. Fitting model parameters to market data

To use either model for contingent claim pricing, the functions $A(0,T)$ and $B(0,T)$ must be estimated. By determining the relationship of $A(0,T)$ and $B(0,T)$ to the initial term structure of interest rates and volatilities, historical data may be used to estimate these functions.

7.5.1. Relationship between $B(0,T)$ and the current term structure of interest rates. We derive the relationship between $B(t,T)$ and the current term structure of interest rates and interest rate volatilities. Here, volatility refers to the standard deviation of proportional, rather than absolute, changes in interest rates.

The time t price of a discount bond with unit maturity value and maturity time T, is simply the unit maturity value discounted to time t using the appropriate rate of interest i.e.

$$\begin{aligned}
P(r,t,T) &= e^{-R(r,t,T)(T-t)} \\
\Rightarrow \quad R(r,t,T) &= -\frac{1}{T-t}\ln P(r,t,T)
\end{aligned}$$

where $R(r,t,T)$ represents the continuously compounded time t rate of interest applicable for the period (t,T). Since, by equation (7.6) the bond price takes the form:

$$P(r,t,T) = A(t,T)e^{-B(t,T)r(t)}$$

we have

$$\ln P(r,t,T) = \ln A(t,T) + \ln e^{-B(t,T)r} = \ln A(t,T) - B(t,T)r(t)$$

and so

$$\begin{aligned} R(r,t,T) &= -\frac{1}{T-t}\left[\ln A(t,T) - B(t,T)r(t)\right] \\ \Rightarrow \quad \frac{\partial R(r,t,T)}{\partial r} &= \frac{B(t,T)}{T-t} \end{aligned} \tag{7.35}$$

Applying Ito's Lemma to determine the stochastic process for $R(r,t,T)$ we have[6]:

$$\begin{aligned} dR &= \frac{\partial R}{\partial t}\,dt + \frac{\partial R}{\partial r}\,dr + \frac{1}{2}\frac{\partial^2 R}{\partial r^2}\,dr dr \\ &= \left(\frac{\partial R}{\partial t} + \frac{\partial R}{\partial r}\,a + \frac{1}{2}\frac{\partial^2 R}{\partial r^2}\,r^2\sigma_r^2(r,t)\right)dt + \frac{\partial R}{\partial r}\,r\sigma_r(r,t)dz \end{aligned}$$

Hence:

$$R(r,t,T)\sigma_R(r,t,T) = r\sigma_r(r,t)\frac{\partial R(r,t,T)}{\partial r} \tag{7.36}$$

where $\sigma_R(r,t,T)$ represents the volatility of $R(r,t,T)$. Now, from (7.35) and (7.36) we have:

$$B(t,T) = \frac{R(r,t,T)\sigma_R(r,t,T)(T-t)}{r\sigma_r(r,t)} \tag{7.37}$$

This equation represents the relationship between $B(t,T)$ and

- the instantaneous short-term interest rate,
- the term structure of spot interest rates,
- instantaneous volatility and
- the term structure of volatilities.

[6] Here the price process of the instantaneous short-term interest rate r is represented as $dr = adt + r\sigma_r(r,t)dz$ where $\sigma_r(r,t)$ is the volatility of r (i.e. standard deviation of relative changes).

Therefore given the current term structure of spot interest rate volatilities, (7.37) may be used to determine $B(0,T)$ for all T.

Alternatively, consider the relationship between spot interest rates and forward rates where $F(r,t,T_1,T_2)$ is the time t forward rate applicable for period (T_1,T_2). We have:

$$e^{R(r,t,T_1)(T_1-t)}e^{F(r,t,T_1,T_2)(T_2-T_1)} = e^{R(r,t,T_2)(T_2-t)}$$

$$\Rightarrow F(r,t,T_1,T_2) = \frac{R(r,t,T_2)(T_2-t) - R(r,t,T_1)(T_1-t)}{T_2-T_1} \tag{7.38}$$

Again, applying Ito's Lemma, allows us to determine the volatility of the forward rate, hence:

$$\begin{aligned} dF &= \frac{\partial F}{\partial t}\,dt + \frac{\partial F}{\partial r}\,dr + \frac{1}{2}\frac{\partial^2 F}{\partial r^2}\,drdr \\ &= \left(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial r}\,a + \frac{1}{2}\frac{\partial^2 F}{\partial r^2}\,r^2\sigma_r^2(r,t)\right)dt + \frac{\partial F}{\partial r}\,r\sigma_r(r,t)dz \end{aligned}$$

and so the standard deviation of $F(r,t,T_1,T_2)$ is:

$$F(r,t,T_1,T_2)\sigma_F(r,t,T_1,T_2) = r\sigma_r(r,t)\frac{\partial F(r,t,T_1,T_2)}{\partial r} \tag{7.39}$$

However, $F(r,t,T_1,T_2)$ is a function of $R(r,t,T_1) \equiv R_1$ and $R(r,t,T_2) \equiv R_2$ which in turn, are functions of r, therefore:

$$\frac{\partial F}{\partial r} = \frac{\partial F}{\partial R_1}\frac{\partial R_1}{\partial r} + \frac{\partial F}{\partial R_2}\frac{\partial R_2}{\partial r}$$

and so by (7.35) and (7.38) we have:

$$\begin{aligned} \frac{\partial F}{\partial r} &= -\frac{T_1-t}{T_2-T_1}\frac{B(r,T_1)}{T_1-t} + \frac{T_2-t}{T_2-T_1}\frac{B(r,T_2)}{T_2-t} \\ &= \frac{B(r,T_2)-B(r,T_1)}{T_2-T_1} \end{aligned}$$

Substituting into (7.39) gives:

$$F(r,t,T_1,T_2)\sigma_F(r,t,T_1,T_2) = r\sigma_r(r,t)\frac{B(r,T_2)-B(r,T_1)}{T_2-T_1}$$

$$\Rightarrow B(r,T_2) - B(r,T_1) = \frac{F(r,t,T_1,T_2)\sigma_F(r,t,T_1,T_2)}{r\sigma_r(r,t)}\,(T_2-T_1) \tag{7.40}$$

The above equation gives the relationship between $B(t,T_1)$ and $B(t,T_2)$ and

- the instantaneous short-term interest rate,
- instantaneous volatility,
- the term structure of forward rates and
- the term structure of forward rate volatilities.

Here (7.37) and (7.40) represent two ways by which $B(0,T)$ may be obtained for all T, either as a function of the current term structure of spot interest rate volatilities or as a function of the current term structure of forward rate volatilities.

7.5.2. Determining $A(0,T)$ from the current term structure. Knowing the current interest rate term structure implies that the current prices of zero coupon bonds are known for all maturities, i.e. we know $P(r,0,T)$ for all T. Evaluating equation (7.18) at $t=0$ gives[7]:

$$P(r,0,T) = A(0,T)e^{-B(0,T)r(0)}$$

Knowing $B(0,T)$ we may determine $A(0,T)$ for all T from the above relationship as

$$A(0,T) = P(r,0,T)e^{B(0,T)r(0)}$$

Alternatively, since it is possible to find analytical solutions to European option prices under the Vasicek model, historical interest rate term structure and option price data can be used to imply the values of $A(0,T)$ and $B(0,T)$ by means of equations (7.20) and (7.24).

7.5.3. Stability of fitted parameters. For a model to be a good description of term structure movements through time, the fitted model parameters $A(t,T)$ and $B(t,T)$ need to remain stable through time. That is, parameters fitted to the term structure of interest rates and interest rate volatilities at time t_1 need to be the same as the parameters fitted to the term structure of interest rates and interest rate volatilities at time T, $t_1 \neq T$. Hence, a model fitted at one time, should correctly describe the term structure at some other time.

The extended Vasicek model does not meet these criteria and hence appears to be unsuitable for this application. However, the goal here has been to develop a model which correctly values most of the interest rate contingent claims in the market. Initially fitting the model to observed prices of vanilla instruments then allows more exotic instruments to be valued consistently.

7.6. Conclusion

Both the Vasicek and CIR models, as discussed in Chapters 1 and 2, incorporate a deterministically mean reverting process. In these models, the mean reversion is incorporated without additional assumptions about the future behaviour of the short-term interest rate volatility. This is a highly desirable feature. By allowing time-dependent parameters and therefore the matching

[7] The explicit functional dependence of r on current time t has, until now, been suppressed to streamline the notation. Here $r(0)$ explicitly denotes the time $t=0$ short-term interest rate.

of any arbitrary initial yield curve, HW manage to overcome one of the major drawbacks of the Vasicek and CIR models. The HW-extended Vasicek model is usually implemented with constant absolute volatility and reversion speed. However, in some yield curve environments, such as rising term structure of rates and declining term structure of volatilities, this version of the model provides a rather poor fit to observed cap prices [**45**]. On the other hand, allowing time-dependent absolute volatility and reversion speed results in unsuitable behaviour of short-term interest rate volatilities in the future. If these shortcomings are recognised and accounted for during the calibration process, the extended Vasicek model can be of great value, since it allows for closed-form solutions of discount bond and discount bond option prices.

CHAPTER 8

The Black, Derman and Toy One-Factor Interest Rate Model

Black, Derman and Toy (BDT) [**6**] make use of a binomial tree approach to model interest rates in a discrete time framework. The model has one fundamental factor, the short-term interest rate, which is used to determine all rates and security prices. The current term structure of interest rates and related volatilities are used to construct a binomial tree of possible short-term interest rates in the future. Since an interest rate sensitive security is characterised by its payoff at expiry, the constructed tree of possible interest rates is used to determine the current price of a security by means of an iterative procedure.

8.1. Model characteristics

The fundamental variable which drives security prices within the model is the short-term interest rate, which is defined as the annualised one period rate of interest.

The model inputs are a set of long-term interest rates of various maturities and their corresponding volatilities. Hence a yield curve and a volatility curve are required to calibrate the model.

These inputs are used to determine mean values and volatilities of future realisation of the short-term interest rate. As the input yield and volatility curves change, so do the means and volatilities of future short-term interest rates. Changes in future volatility have an impact on the degree of mean reversion.

As with most models, the assumption of perfect markets is made, hence:

- changes in the yields of all zero coupon bonds are perfectly correlated,
- the expected one period returns are the same for all securities,
- short-term interest rates are lognormally distributed and
- the market is free of taxes and transaction costs.

The lognormality feature holds several advantages for calibration of the model [**45**]. Negative interest rates are prevented and the volatility input may be specified in percentage terms, i.e. the volatility refers to relative price

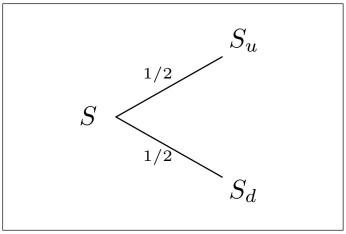

FIGURE 8.1. Tree with one time step

moves. This is the market convention for quoting volatilities, so calibration to market-observed volatilities is simplified.

8.2. Pricing contingent claims

The short-term interest rate at each node in the tree is found such that the term structure produced by the model matches the current observed term structure. European-style contingent claims may then be priced. The value at a node is the discounted expected value one time period in the future. Since the binomial tree is calibrated to the market-observed risk-free rate, the contingent claim is priced in a risk-neutral environment, where the probabilities of an up and down move are equal. Hence our expectation of the price of the contingent claim after one period is:

$$\tfrac{1}{2}\left(S_u + S_d\right)$$

where S_u and S_d are the prices of the contingent claim after an up and down move respectively. Discounting by the current one period interest rate r, the current price of the contingent claim S, is:

$$S = \frac{\frac{1}{2}\left(S_u + S_d\right)}{1+r} \tag{8.1}$$

This method may be used to determine the price at any node in the tree, from the prices one step in the future. Iterative application of (8.1) allows valuation of contingent claims of any duration, as long as the tree of short-term interest rates extends sufficiently far into the future.

To value bond options, the tree must first be used to determine the bond value at every node according to the interest rate associated with that node.

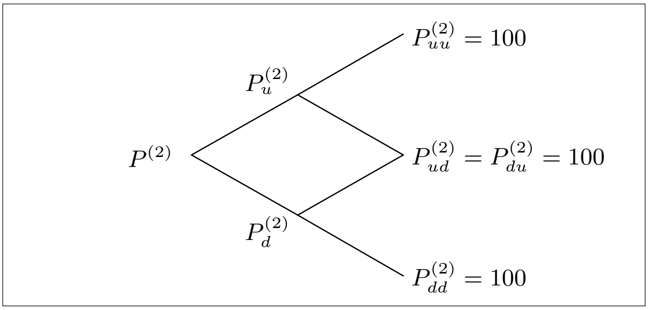

FIGURE 8.2. Valuation of a 2-year zero coupon bond

Then, making use of the known option value at expiry[1], and working backwards through time, the bond option value may be found at every node prior to expiry.

8.3. Calibrating the lattice to an observed term structure

The market convention for quoting a term structure is in terms of annualised yields of zero coupon bonds of various maturities. Therefore if y is the N year rate, then the current price of the associated zero coupon bond[2] is:

$$P = \frac{100}{(1+y)^N}$$

Calibration of the tree involves finding the one-period yield (short-term interest rate) at each node, such that the observed term structure is matched. The following methodology for calibrating the binomial tree to a market-observed interest rate and volatility term structure is outlined by BDT [**6**].

Consider the interest rate and volatility term structures observed at time $t = 0$, represented by $\{(y_i, \sigma_i) : i = 1, \ldots, N\}$ where y_i is the yield of a zero coupon bond with maturity $t = i$ and σ_i the corresponding volatility. For simplicity, assume that each time interval in the tree is one year and we wish to calibrate the tree such that the yield and volatility term structures are matched. For shorter time steps we would require the observed yields and volatilities at more frequent time intervals.

At time $t = 0$, the short-term interest rate may be taken directly from the observed term structure as the 1-year yield, hence $r_0 = y_1$. To determine the short-term interest rates at time $t = 1$, make use of the observed 2-year yield

---

[1]This is the option payoff at expiry, hence for a call option the payoff is spot less the strike, while for a put option it is the strike less the spot.

[2]Assume all zero coupon bonds have a maturity value of 100.

and associated volatility. The current value of a 2-year zero coupon bond, $P_0^{(2)} = P^{(2)}$ is calculated as:

$$P^{(2)} = \frac{100}{(1+y_2)^2} \tag{8.2}$$

At time $t = 1$, the 2-year zero coupon bond has one year left to run, so its prices, $P_u^{(2)}$ and $P_d^{(2)}$ (where the subscript indicates an up or down move in the short-term interest rate), may be found as:

$$P_u^{(2)} = \frac{100}{1+r_u} \qquad P_d^{(2)} = \frac{100}{1+r_d} \tag{8.3}$$

where r_u and r_d are the time $t = 1$ short-term interest rates resulting from an up move and down move respectively. These time $t = 1$ bond prices must be such that discounting by the $t = 0$ short-term interest rate r_0, we obtain the time $t = 0$ price of the 2-year zero coupon bond, $P^{(2)}$. That is:

$$P^{(2)} = \frac{\frac{1}{2}P_u^{(2)} + \frac{1}{2}P_d^{(2)}}{(1+r_0)} \tag{8.4}$$

We also need to match the term structure of volatilities. The standard deviation of the short-term interest rate at time $t = 1$ is matched to the volatility of the 2-year yield σ_2, hence[3]:

$$\sigma_2 = \frac{\ln r_u - \ln r_d}{2} \tag{8.5}$$

The equations (8.3)–(8.5) are solved simultaneously for the four unknowns:

$$\begin{aligned}
P_u^{(2)}(1+r_u) &= 100 && \text{(8.6a)}\\
P_d^{(2)}(1+r_d) &= 100 && \text{(8.6b)}\\
P_u^{(2)} + P_d^{(2)} &= 2P^{(2)}(1+r_0) && \text{(8.6c)}\\
r_u &= r_d\, e^{2\sigma_2} && \text{(8.6d)}
\end{aligned}$$

[3]Consider a random variable X. At time $t = t^*$, X may take on two possible values, x_1 and x_2, each with probability $\frac{1}{2}$. Without loss of generality, let $x_1 \geq x_2$. Hence:

$$\begin{aligned}
\operatorname{var}(X) &= \mathbb{E}\left[X^2\right] - \mathbb{E}\left[X\right]^2\\
&= \frac{x_1^2}{2} + \frac{x_2^2}{2} - \left(\frac{x_1+x_2}{2}\right)^2\\
&= \left(\frac{x_1-x_2}{2}\right)^2\\
\Rightarrow \operatorname{stddev}(X) &= \frac{x_1-x_2}{2}
\end{aligned}$$

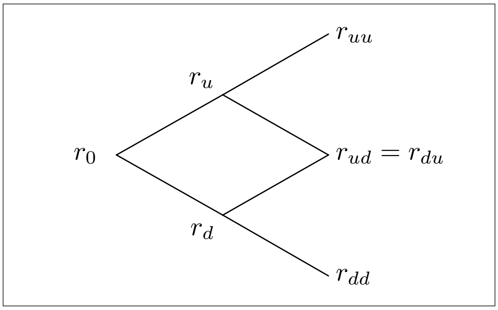

FIGURE 8.3. Tree of short-term interest rates out to 2 years

The resulting $t = 1$ short-term interest rates r_u and r_d are exactly consistent with the 2-year term structure of yields and volatilities. To determine the possible short-term interest rates at $t = 2$, make use of the 3-year yield and volatility. There are three possible short-term interest rates at $t = 2$, r_{uu}, $r_{du} = r_{ud}$ and r_{dd}, but only two known values, the yield and volatility, to which we calibrate. This means that the solution is not unique. However, the short-term interest rate volatility is a function of time only, therefore:

$$\begin{aligned} \frac{\ln r_{uu} - \ln r_{dd}}{2} &= \frac{\ln r_{ud} - \ln r_{dd}}{2} \\ \frac{r_{uu}}{r_{ud}} &= \frac{r_{ud}}{r_{dd}} \\ \Rightarrow r_{ud}^2 &= r_{uu} r_{ud} \end{aligned}$$

and hence, we need only match two short-term interest rates to two observed values and can find a unique solution.

8.4. Continuous time equivalent

The original specification of the BDT model is in a discrete time framework. There are several disadvantages associated with this type of formulation. The algorithmic manner in which one must view the model makes it difficult to identify the embedded assumptions and their implications, affecting, for example, the characteristics of the mean reversion. While the lognormality of the short-term interest rate is a positive feature since it precludes negative interest rates, it does make analytical analysis cumbersome [**45**]. We examine the continuous time equivalent of the BDT model[4]. Consider a process for the short-term interest rate $r(\cdot)$, as follows[5]:

$$r(t) = u(t) \exp\left(\sigma(t) z(t)\right) \tag{8.7}$$

where

$u(t)$	–	time t median of the short-term interest rate distribution,
$\sigma(t)$	–	short-term interest rate volatility at time t,
$z(t)$	–	standard Brownian motion.

To examine the nature of the stochastic process driving the short-term interest rate as modelled by BDT, we must examine the evolution of $\ln r(t)$ where

$$\begin{aligned} d\ln r(t) &= \frac{\partial \ln r(t)}{\partial r}\, dr + \frac{1}{2}\frac{\partial^2 \ln r(t)}{\partial r^2}\, drdr \\ &= \tfrac{1}{r}\, dr - \tfrac{1}{2r^2}\, drdr \end{aligned} \tag{8.8}$$

[4]Rebonato [**45**] details an analysis of binomial pricing within a lattice, which leads to the derivation of the continuous equivalent of the BDT model.

[5]In the BDT model, a lognormal distribution of the short-term interest rate is assumed. This implies that $\ln r(t)$ is normally distributed. At each time step t in a binomial lattice, we have $t+1$ possible states of the world and hence $t+1$ possible values of the one period rate. Consider time $t = 1$: we have two possible states of the world and interest rates denoted $r(1,1)$ and $r(1,-1)$. The mean short-term interest rate at this time, $\ln r_m(1)$ may be calculated as:

$$\ln r_m(1) = \tfrac{1}{2}\left[\ln r(1,1) + \ln r(1,-1)\right]$$

Knowing the time $t = 0$ percentage volatility of the short-term interest rate, i.e. $\sigma(0)$, the standard deviation of the time $t = 1$ short-term interest rates is represented as:

$$\sigma(0)\sqrt{\Delta t} = \frac{\ln r(1,1) - \ln r(1,-1)}{2} \quad \Rightarrow \quad r(1,1) = r(1,-1)\exp\left[2\sigma(0)\sqrt{\Delta t}\right]$$

and hence, we find each of the two possible rates as an offset from the median rate of interest, $r_m(1)$:

$$\begin{aligned} r(1,1) &= r_m(1)\exp\left[\sigma(0)\sqrt{\Delta t}\right] \\ \text{and} \quad r(1,-1) &= r_m(1)\exp\left[-\sigma(0)\sqrt{\Delta t}\right] \end{aligned}$$

Note that the mean of the distribution of the logarithm of the short-term interest rate corresponds to the median of the lognormal distribution of the short-term interest rate. $\ln r(t)$ is normally distributed with mean $\ln r_m(1)$ and $r(t)$ is lognormally distributed with median $r_m(1)$.

Since $r(t) = r(t, z(t))$ Ito's Lemma gives:

$$dr = \frac{\partial r}{\partial t}\,dt + \frac{\partial r}{\partial z}\,dz + \frac{1}{2}\frac{\partial^2 r}{\partial z^2}\,dzdz \tag{8.9}$$

where

$$\begin{aligned}
\frac{\partial r}{\partial t}\,dt &= \frac{\partial u(t)}{\partial t}\ \exp\left(\sigma(t)z(t)\right)dt \\
&\quad + u(t)\frac{\partial \sigma(t)z(t)}{\partial t}\ \exp\left(\sigma(t)z(t)\right)dt \\
&= \frac{\partial u(t)}{\partial t}\ \exp\left(\sigma(t)z(t)\right)dt \\
&\quad + u(t)\left[z(t)\frac{\partial \sigma(t)}{\partial t}\right]\exp\left(\sigma(t)z(t)\right)dt
\end{aligned} \tag{8.10}$$

$$\frac{\partial r}{\partial z}\,dz = u(t)\sigma(t)\exp\left(\sigma(t)z(t)\right)dz \tag{8.11}$$

$$\frac{\partial^2 r}{\partial z^2}\,dzdz = u(t)\sigma^2(t)\exp\left(\sigma(t)z(t)\right)dt \tag{8.12}$$

Substituting (8.7), (8.9)–(8.12) into (8.8) we have[6]:

$$\begin{aligned}
d\ln r(t) &= \frac{1}{u(t)\exp\left(\sigma(t)z(t)\right)} \\
&\quad \times \left(\frac{\partial u(t)}{\partial t}\ \exp\left(\sigma(t)z(t)\right)dt + u(t)\left[z(t)\frac{\partial \sigma(t)}{\partial t}\right]\exp\left(\sigma(t)z(t)\right)dt\right. \\
&\quad \left. + u(t)\sigma(t)\exp\left(\sigma(t)z(t)\right)dz + \tfrac{1}{2}\,u(t)\sigma^2(t)\exp\left(\sigma(t)z(t)\right)dt\right) \\
&\quad - \frac{1}{2u^2(t)\exp\left(2\sigma(t)z(t)\right)}\ u^2(t)\sigma^2(t)\exp\left(2\sigma(t)z(t)\right)dt \\
&= \frac{\partial \ln u(t)}{\partial t}\,dt + z(t)\frac{\partial \sigma(t)}{\partial t}\,dt + \sigma(t)dz + \tfrac{1}{2}\,\sigma^2(t)dt - \tfrac{1}{2}\,\sigma^2(t)dt \\
&= \frac{\partial \ln u(t)}{\partial t}\,dt + \frac{\ln r(t) - \ln u(t)}{\sigma(t)}\frac{\partial \sigma(t)}{\partial t}\,dt + \sigma(t)dz \\
&= \left[\frac{\partial \ln u(t)}{\partial t} - \frac{\partial \ln \sigma(t)}{\partial t}\left(\ln u(t) - \ln r(t)\right)\right]dt + \sigma(t)dz
\end{aligned} \tag{8.13}$$

---

[6]From (8.7) we have:

$$\begin{aligned}
r(t) &= u(t)\exp\left(\sigma(t)z(t)\right) \\
\Rightarrow \quad z(t) &= \frac{\ln r(t) - \ln u(t)}{\sigma(t)}
\end{aligned}$$

This is a mean reverting process that is explicitly dependent on the median of the distribution. This median is implicitly determined during the tree-fitting procedure.

If we allow the volatility to be constant, i.e. $\sigma(t) \equiv \sigma$, then $\frac{\partial \ln \sigma(t)}{\partial t} = 0$ and there is no mean reversion. The logarithm of the short-term interest rate follows a simple diffusion with a drift that follows the logarithm of the median of the distribution. That is, the process simplifies to:

$$d \ln r(t) = \frac{\partial \ln u(t)}{\partial t} dt + \sigma dz \tag{8.14}$$

For a volatility that decays with time, i.e. $\frac{\partial \ln \sigma(t)}{\partial t} < 0$, the reversion speed becomes positive and the logarithm of the short-term interest rate $\ln r(t)$ reverts to the logarithm of the median $\ln u(t)$.

This assumption of a decaying short-term interest rate volatility is necessary to ensure that the unconditional variance of the short-term interest rate $\sigma^2(t)t$, does not increase without bound as t increases and the mean-reverting nature of the short-term interest rate process is maintained.

The reversion speed determines the volatility of rates for various maturities. In this model the reversion speed is a unique function of the short-term interest rate volatility $\sigma(t)$, and hence we conclude that the entire term structure of volatilities is fully determined by the future short-term interest rate volatilities. This is unique to the BDT model, since other models, e.g. the Hull–White model [**28**], specify the reversion speed as an independent parameter. This artificial link created between the future short-term interest rate volatilities and the term structure of volatilities may be seen as one of the shortcomings of this model.

8.5. A fundamental flaw

Let us revisit the calibration methodology outlined in §8.3. It is easy to see that this methodology contains a fundamental flaw. No distinction is made between the current term structure of interest rate volatilities and future volatilities of the short-term interest rate. Equation (8.5) shows that the standard deviation of the time $t = 1$ short-term interest rate is matched to the volatility of the 2-year yield, and hence the volatility of the short-term interest rate at time $t = 1$ is matched to the current volatility of the 2-year yield. Assuming time steps of 1 year, this may be generalised as follows: the time $t = t^*$ volatility of the short-term interest rate is equated to the current (time $t = 0$) volatility of the $(t^* + 1)$-year yield. Surprisingly, there are few texts that make mention of this flaw.

The artificial link between the future short-term interest rate volatilities and the term structure of volatilities is often cited as a major drawback of the BDT model. In fact Rebonato [**45**] discusses the fact that "the term structure of volatilities is completely determined by the specification of the future volatility of the short rate" and examines the shape of the term structure of volatilities for various functional forms of the short-term interest rate volatility. However, he does not explicitly show the relationship between the volatility term structure and the short-term interest rate volatility function, nor does he explicitly discuss the calibration to market-observed volatilities thereby side-stepping the issue.

A slightly more complex procedure to that outlined in §8.3 is required to build a binomial tree representing both the observed interest rate and volatility term structures. The following methodology [**15**] allows matching of the observed interest rate and volatility term structures.

8.5.1. Preliminaries. We begin with the observed interest rate and volatility term structures represented by: $P(i)$ as the price of a discount bond[7] maturing at time $i\Delta t$ and $\sigma_R(i)$ as the volatility of the yield on this bond. Let

Δt	–	time step size chosen for the tree,
$u(i)$	–	median short-term interest rate at time $i\Delta t$,
$\sigma(i)$	–	volatility of the short-term interest rate at time $i\Delta t$,
$r_{i,j}$	–	short-term interest rate at time $i\Delta t$ node j, applicable for the period $[i\Delta t, (i+1)\Delta t]$,
$d_{i,j}$	–	the time $i\Delta t$, state j value of a discount bond maturing at time $(i+1)\Delta t$, hence: $d_{i,j} = 1/(1 + r_{i,j}\Delta t)$. In [**15**] this is called the one-period discount factor at node (i,j).
$Q_{i,j}$	–	time 0 value of a security paying: 1 if node (i,j) is reached, 0 otherwise.

The $Q_{i,j}$s are in fact Arrow–Debreu securities. For a discussion on this topic see [**45**]. They may be viewed as discounted probabilities; hence, by definition $Q_{0,0} = 1$.

Hence, the current (time $t = 0$) price of a discount bond maturing at time $(i+1)\Delta t$ may be expressed as:

$$P(i+1) = \sum_j Q_{i,j}\, d_{i,j} \tag{8.15}$$

[7]This discount bond price is related to yield $R(i)$ by:

$$P(i) = \exp\left(-R(i)i\Delta t\right)$$

for $j \in \mathcal{A}$, where $\mathcal{A}$ is the set of all possible states at time i. The calibration of the tree is by forward induction, where the time $i\Delta t$ Arrow–Debreu securities are updated using the already known Arrow–Debreu securities at time $(i-1)\Delta t$ as follows:

$$\begin{aligned} Q_{i,i} &= \tfrac{1}{2}\, Q_{i-1,i-1}\, d_{i-1,i-1} \\ Q_{i,j} &= \tfrac{1}{2}\, Q_{i-1,j-1}\, d_{i-1,j-1} + \tfrac{1}{2}\, Q_{i-1,j+1}\, d_{i-1,j+1} \\ Q_{i,-i} &= \tfrac{1}{2}\, Q_{i-1,-i+1}\, d_{i-1,-i+1} \end{aligned} \tag{8.16}$$

8.5.2. Fitting the interest rate and volatility term structures. From (8.7) the short-term interest rate at each node (i, j) may be represented as:

$$r_{i,j} = u(i) \exp\left(\sigma(i) j \sqrt{\Delta t}\right) \tag{8.17}$$

First we define the notation used to index the nodes of the tree. At starting time $i = 0$, there is a single state $j = 0$. At each subsequent time i there are $(i + 1)$ possible states indexed as $j = -i, -i + 2, \ldots, i - 2, i$. The state index represents the net moves required to reach it. For example, at time $i = 4$, there are 5 possible states with indices $\{-4; -2; 0; 2; 4\}$ corresponding to 4 down moves, 1 up and 3 down moves (net 2 down moves), 2 up and 2 down moves (net zero move), 3 up and 1 down move (net 2 up moves) and 4 up moves respectively.

Now from the initial node $(0, 0)$ at the root of the tree, we have a possible up move and possible down move. Hence node $(1, 1)$ is denoted U and node $(1, -1)$ is denoted D. At these nodes define the following:

P_U^i, P_D^i	–	price, at nodes U and D respectively, of discount bond with maturity $i\Delta t$ after an initial up or down move, $i \geq 1$,
R_U^i, R_D^i	–	the discount bond yields at nodes U and D respectively, corresponding to the above discount bond prices,

These values of $P_U(i)$ and $P_D(i)$ must be consistent with the observed values of $P(i)$ and $\sigma_R(i)$. Therefore the following relationships must hold:

$$\frac{1}{1 + r_{0,0}\Delta t} \left(\tfrac{1}{2} P_U(i) + \tfrac{1}{2} P_D(i)\right) = P(i) \qquad i = 2, \ldots, N \tag{8.18a}$$

$$\sigma_R(i)\sqrt{\Delta t} = \tfrac{1}{2} \ln \frac{\ln P_U(i)}{\ln P_D(i)} \qquad i = 2, \ldots, N \tag{8.18b}$$

Equations (8.18) may be solved simultaneously for $P_U(i)$ and $P_D(i)$ as:

$$P_D(i)^{\exp(2\sigma_R(i)\sqrt{\Delta t})} + P_D(i) = 2P(i)\big(1 + r_{0,0}\Delta t\big) \tag{8.19}$$

and

$$P_U(i) = P_D(i)^{\exp(2\sigma_R(i)\sqrt{\Delta t})} \tag{8.20}$$

where (8.19) must be solved numerically for $P_D(i)$.

As in §8.5.1 we define state prices, this time corresponding to nodes U and D, as follows:

$Q_{U,\,i,\,j}$ – value at node U of security paying 1 if node (i,j) is reached, 0 otherwise,

$Q_{D,\,i,\,j}$ – value at node D of security paying 1 if node (i,j) is reached, 0 otherwise.

By definition $Q_{U,\,1,\,1} = 1$ and $Q_{U,\,1,\,-1} = 1$ and so from (8.15) the values at nodes U and D of a discount bond maturing at time $(i+1)\Delta t$ may be written as[8]:

$$P_U(i+1) = \sum_j Q_{U,\,i,\,j}\, d_{i,j} \qquad i = 1, \ldots, N-1 \tag{8.21}$$

$$P_D(i+1) = \sum_j Q_{D,\,i,\,j}\, d_{i,j} \qquad j \in \{-i; -i+2; \ldots; i-2; i\} \tag{8.22}$$

respectively.

The state prices may then be updated as in (8.16):

$$Q_{U,\,i,\,i} = \tfrac{1}{2}\, Q_{U,\,i\,-\,1,\,i\,-\,1}\, d_{i-1,i-1} \tag{8.23a}$$

$$Q_{U,\,i,\,j} = \tfrac{1}{2}\, Q_{U,\,i\,-\,1,\,j\,-\,1}\, d_{i-1,j-1} + \tfrac{1}{2}\, Q_{U,\,i\,-\,1,\,j\,+\,1}\, d_{i-1,j+1} \tag{8.23b}$$

$$Q_{U,\,i,\,-i\,+\,2} = \tfrac{1}{2}\, Q_{U,\,i\,-\,1,\,-i\,+\,3}\, d_{i-1,-i+3} \tag{8.23c}$$

and

$$Q_{D,\,i,\,-i} = \tfrac{1}{2}\, Q_{D,\,i\,-\,1,\,-i\,+\,1}\, d_{i-1,-i+1} \tag{8.23d}$$

$$Q_{D,\,i,\,j} = \tfrac{1}{2}\, Q_{D,\,i\,-\,1,\,j\,-\,1}\, d_{i-1,j-1} + \tfrac{1}{2}\, Q_{D,\,i\,-\,1,\,,\,j\,+\,1}\, d_{i-1,j+1} \tag{8.23e}$$

$$Q_{D,\,i,\,i\,-\,2} = \tfrac{1}{2}\, Q_{D,\,i\,-\,1,\,i\,-\,3}\, d_{i-1,i-3} \tag{8.23f}$$

[8]Here N is the total number of time steps in the tree.

8.5.3. Basic algorithm. First, set the initial values:

$$\begin{aligned}
r_{0,0} = u(0) &= \frac{e^{(R(1)\Delta t)} - 1}{\Delta t} \\
Q_{U,1,1} &= 1 \\
Q_{D,1,-1} &= 1 \\
\sigma(0) &= \sigma_R(1) \\
d_{0,0} &= \frac{1}{1 + r_{0,0}\Delta t}
\end{aligned}$$

Now for each $i = 1, \ldots, N-1$

(1) Using a numerical method such as Newton–Raphson solve (8.19) for $P_D(i+1)$ and then solve for $P_U(i+1)$ by means of (8.20).

(2) Now making use of $P_D(i+1)$ and $P_U(i+1)$ derived above, equations (8.21) and (8.22) may be used to solve for $\sigma(i)$ and $u(i)$. The following substitution is made:

$$d_{i,j} = \frac{1}{1 + u(i)\exp\left(\sigma(i) j\sqrt{\Delta t}\right)}$$

Hence:

$$P_U(i+1) = \sum_j \frac{Q_{U,i,j}}{1 + u(i)\exp\left(\sigma(i) j\sqrt{\Delta t}\right)} \qquad i = 1, \ldots, N-1$$

$$P_D(i+1) = \sum_j \frac{Q_{D,i,j}}{1 + u(i)\exp\left(\sigma(i) j\sqrt{\Delta t}\right)} \qquad j \in \{-i; -i+2; \ldots; i-2; i\}$$

and $\sigma(i)$ and $u(i)$ may be found using a two dimensional Newton–Raphson (or other optimisation) technique.

(3) Using these calculated values of $\sigma(i)$ and $u(i)$ the one period short-term interest rates and discount factors may be found for each node $j = -i, \ldots, i$ using (8.17) and the definition of the discount factor as $d_{i,j} = 1/(1 + r_{i,j}\Delta t)$.

(4) $Q_{U,i,j}$ and $Q_{D,i,j}$ may now be updated using equations (8.23).

The fundamental concept behind this calibration methodology is rather simple. The observed term structure of volatilities represents the volatilities of *current* rates of various maturities. A shift up and shift down of the entire term structure is simulated by the U and D nodes. The magnitude of this up and down shift is determined by the bond yield volatility (as represented by (8.18b)). This allows the volatility term structure information to be incorporated. Comparing this to the original BDT calibration methodology in §8.3 highlights their error of matching the initial time volatility of the i-year interest rate to the short-term interest rate volatility at year $(i-1)$. Their methodology could be used to build a tree representing the interest rate term structure and future volatilities of the short-term interest rate. Obviously this is only possible if one has a view of short-term interest rate volatilities at each time interval in the future.

8.6. Conclusion

The BDT model has several positive features:

- for positive value of the decay factor (reversion speed) the general shape of the term structure of volatilities which is captured within the BDT model is consistent with the market-observed volatility term structure.
- due to the lognormal process assumed for the short-term interest rate, calibration to market prices becomes much simpler. It is possible to fit the model to both the yield curve and to cap volatilities at the same time. Hence the model can simultaneously reproduce the prices of various maturity caps, displaying a declining term structure of volatilities.

However, it also displays several problems:

- as with all one factor interest rate models the changes in rates of various maturities are by and large parallel which is not consistent with market observation. Hence the BDT model is not able to capture a tilting effect on the yield curve. This would require a second factor.
- no specification is made of the evolution, through time, of the term structure of volatilities.
- since the future short-term interest rate volatilities fully determine the term structure of volatility it is impossible to specify one independently of the other.

This model was developed by practitioners for practitioners and hence allows for easy calibration to observed data and easy pricing of European and American style contingent claims.

CHAPTER 9

The Black and Karasinski Model

The discrete time Black, Derman and Toy model [**6**], discussed in Chapter 8, makes provision for two time-dependent factors: the mean short-term interest rate and the short-term interest rate volatility. The continuous time equivalent of the model clearly shows that the rate of mean reversion is a function of the volatility. This is equivalent to future short-term interest rate volatilities being fully determined by the observed volatility term structure. This dependence makes it impossible to specify these two factors independently.

Black and Karasinski (BK) [**7**] develop a model, within a discrete time framework, where the target rate, mean reversion rate and local volatility are deterministic functions of time. The specification of three time-dependent factors allows 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 horizon. The standard assumptions underlying perfect markets are also made.

9.1. The lognormality assumption

Ideally one wants a process for the short-term interest rate such that negative interest rates are prevented, but the zero level may be reached and maintained for extended periods of time. None of the processes examined thus far, i.e. normal, lognormal and square root processes, satisfy both these requirements. A lognormal process does not admit a zero interest rate, while the square root process makes the zero level a reflecting barrier.

BK use a lognormal process. A lognormal distribution is fully described by its mean and variance, which are functions of time, so we have a different lognormal distribution of the short-term interest rate at each future time. When mean reversion is combined with a lognormal model, we have three time-dependent factors, an example being the BDT model:

$$d(\ln r) = (\theta(t) - \phi(t)\ln r)\,dt + \sigma(t)dz$$

However, here $\phi(t)$ is a function of $\sigma(t)$. Dropping this functional dependence, and letting $\mu(t)$ be the target interest rate, i.e. the reversion level, the BK model may be written as:

$$d(\ln r) = \phi(t)\left(\ln \mu(t) - \ln r\right) dt + \sigma(t) dz$$

where $\phi(t)$ is the speed of the mean reversion and $\sigma(t)$ the local volatility, i.e. the volatility of the short-term interest rate. BK calibrate their model to the initial observed interest rate and volatility term structures as well as the observed cap curve. The cap curve gives the prices of at-the-money caps, which pay the difference between the forward rate (strike) and the realised short-term interest rate at maturity. BK do not attempt to specify a process which accurately depicts the evolution of the short-term interest rate, but rather a short-term interest rate process which can be fitted to observed market prices and hence used to price securities in a consistent manner. The future risk-neutral distribution of the short-term interest rate generated by the model is not the true distribution, but rather a distribution which leads to correct option prices.

9.2. Specification of the binomial tree

BDT make use of a binomial tree to specify their lognormal model. Within the binomial tree they are able to match two inputs: the interest rate and volatility term structures. This is done using the location and spacing of the nodes at each time point.

To match three input values, one could use a trinomial tree. However, to avoid the additional complexity of a trinomial tree, BK approach this problem with a binomial tree, but vary the time spacing during its life. This introduces another degree of freedom, allowing all three inputs to be matched. The computational simplicity of a binomial tree is maintained and the risk-neutral probabilities are $\frac{1}{2}$.

9.2.1. Known model inputs. If the input functions defining the model, i.e. $\mu(t)$, $\phi(t)$ and $\sigma(t)$ are known, the binomial tree of short-term interest rates is constructed so as to match these values at each time step. The tree has the following specifications:

- At each time, the (vertical) spacing of the nodes must match the local volatility (volatility of the short-term interest rate). Since volatility is a function of time only, spacings for a given time are equal.
- The drift of the nodes from one time to the next is determined by the target rate.
- The time (horizontal) spacing differs over the life of the binomial tree. This time spacing is calibrated to the mean reversion speed.

Define the following variables:

$$\begin{aligned} \tau_n = t_{n+1} - t_n \quad &- \quad \text{time period between two consecutive time nodes,} \\ \phi_n = \phi(t_n) \quad &- \quad \text{mean reversion speed at time } t = n, \\ \sigma_n = \sigma(t_n) \quad &- \quad \text{local volatility at time } t = n. \end{aligned}$$

Mean reversion is defined as the speed with which the short-term interest rate tends towards the target rate. As the short-term interest rate gets closer to this target rate, the local volatility decreases. Hence, the mean reversion may be equated to the rate of change of local volatility, which is represented by:

$$\phi_n = \frac{1}{\tau_n}\left(1 - \frac{\sigma_n\sqrt{\tau_n}}{\sigma_{n-1}\sqrt{\tau_{n-1}}}\right) \tag{9.1}$$

For positive mean reversion $\sigma_n\sqrt{\tau_n} < \sigma_{n-1}\sqrt{\tau_{n-1}}$ and $1-\frac{\sigma_n\sqrt{\tau_n}}{\sigma_{n-1}\sqrt{\tau_{n-1}}}$ gives the percentage decrease in volatility from time period τ_{n-1} to τ_n. From (9.1) we may find τ_n as a function of τ_{n-1} and the speed of mean reversion. Hence, at each time node, we may determine the size of the next time step, dependent on the speed of mean reversion. Using equation (9.1) we have:

$$\begin{aligned} \phi_n &= \frac{1}{\tau_n} - \frac{\sigma_n}{\sigma_{n-1}\sqrt{\tau_n}\sqrt{\tau_{n-1}}} \\ \Rightarrow 0 &= \frac{1}{\tau_n} - \frac{1}{\sqrt{\tau_n}}\left(\frac{\sigma_n}{\sigma_{n-1}}\right)\frac{1}{\sqrt{\tau_{n-1}}} - \phi_n \end{aligned} \tag{9.2}$$

which is a quadratic polynomial in $\frac{1}{\sqrt{\tau_n}}$ with roots:

$$\begin{aligned} \frac{1}{\sqrt{\tau_n}} &= \frac{1}{2}\left(\left(\frac{\sigma_n}{\sigma_{n-1}}\right)\frac{1}{\sqrt{\tau_{n-1}}} \pm \sqrt{\left(\frac{\sigma_n}{\sigma_{n-1}}\right)^2\frac{1}{\tau_{n-1}} + 4\phi_n}\right) \\ &= \frac{1 \pm \sqrt{1 + 4\phi_n\left(\sigma_{n-1}/\sigma_n\right)^2\tau_{n-1}}}{2\sqrt{\tau_{n-1}}\left(\sigma_{n-1}/\sigma_n\right)} \end{aligned}$$

Since by definition $\frac{1}{\sqrt{\tau_n}} > 0$, only one of the roots is an admissible solution to (9.2) and:

$$\tau_n = \tau_{n-1}\left(\frac{4\left(\sigma_{n-1}/\sigma_n\right)^2}{\left(1 + \sqrt{1 + 4\phi_n\left(\sigma_{n-1}/\sigma_n\right)^2\tau_{n-1}}\right)^2}\right) \tag{9.3}$$

Hence the time spacing is dynamically constructed, with the next time step size determined at each node after all its associated variables σ_n, ϕ_n etc.

have been determined. The initial time step τ_0 is chosen according to the required accuracy. Small τ_0 equates to very fine time spacing, which produces more accurate results. For positive mean reversion speed, the time spacing decreases through time, and the higher the speed of mean reversion the more pronounced this decrease.

9.2.2. Known model outputs. If we already have the model outputs, that is the interest rate, volatility and cap term structures, we need to find the corresponding values of the model inputs, $\mu(t)$, $\phi(t)$ and $\sigma(t)$. Here, time is divided into segments which are subdivided into time steps. The values of μ, ϕ and σ, applicable for the first time segment, are chosen so as to match the outputs at the end of this segment. Similarly μ, ϕ and σ applicable during the second time segment are chosen such that the outputs are matched at the end of this segment. Using this methodology, we find the implied target rate, mean reversion speed and short-term interest rate volatility for each time segment. These implied values do not specify the real-world evolution of the short-term interest rate, but they do specify a short-term interest rate process in a one-factor world which produces the required security prices.

When the model outputs change, that is the interest rate and volatility term structures observed in the market shift, the tree needs to be recalibrated to determine new parameters of the implied process.

Ideally, we would like to determine a general interest rate process as a function of several parameters. Re-estimation of the process should yield the same parameters. This type of model would be a true description of the interest rate process and could be used to give valuations at any time.

9.3. Matching the lognormal distribution

We have assumed the short-term interest rate has a lognormal distribution at any time horizon. This means we require only a mean and standard deviation to fully specify its distribution. However, in the BK model, three factors are required to describe the short-term interest rate process – the target rate, mean reversion speed and local volatility. This means that for a given time horizon, the solution is not unique and the distribution of short-term interest rates may be matched by a family of possible processes.

These processes will differ in their mean reversion and local volatility characteristics. Strong mean reversion means a move away from the target rate is quickly reversed, which is not the case for weaker mean reversion. Hence, a narrow (wide) distribution of the short-term interest rate in the future may result from either strong (weak) mean reversion or low (high) local volatility.

9.4. Conclusion

In a simple and concise extension of the BDT model, BK are able to eliminate one of its most frequently cited shortcomings – the direct but artificial link between the current volatility term structure and future values of short-term interest rate volatility. BK introduce a third time-dependent variable, reversion speed, which allows an additional degree of freedom. Now the interest rate and volatility term structures as well as cap prices can be included in the calibration procedure.

Empirical results of calibration exercises for this model are not widely available. However, based on results of other models[1] attempting to include all three term structures (interest rate, volatility and cap prices), one could suspect an over-parameterisation may result, with future volatility term structures taking on unreasonable shapes.

[1]For example the extended-Vasicek Hull White model discussed in Chapter 7.

CHAPTER 10

The Ho and Lee Model

Models studied in the previous chapters specify the movement of the short-term interest rate and thereby endogenously determine the form of term structure (including its initial value). Ho and Lee (HL) [**27**] developed a model which takes as input, the initial interest rate term structure and derives its subsequent stochastic evolution. Hence the theoretical zero coupon bond prices (that is, those produced by the model) will be exactly consistent with those observed in the market.

HL use all information within the current observed term structure to price contingent claims in such as way as to ensure that profitable arbitrage is precluded.

10.1. Assumptions

The assumptions made by HL are the standard assumptions for a perfect capital market in a discrete time framework.

Assumption 1. The market is frictionless, i.e. there are no taxes or transaction costs and securities are perfectly divisible.

Assumption 2. In a discrete time framework each time period is taken to be one unit of time. Hence a zero coupon bond with term to maturity T pays \$1 at the end of the T^{th} time period (taken from valuation time).

Assumption 3. The bond market is complete, with a bond maturing at the end of each time period n, $n = 0, 1, 2, \ldots$.

Assumption 4. At each time period n, there are a finite number of possible states of the world. At time n, state i, denote the equilibrium price of a T-maturity zero coupon bond as $P_i^{(n)}(T)$. This function is termed a discount function. At any time n, state i, the interest rate term structure is fully described by a series of discount functions.

By its definition as a discount function, $P_i^{(n)}(\cdot)$ must satisfy certain conditions. That is:

$$P_i^{(n)}(T) \geq 0 \quad \forall\, T, i \text{ and } n \tag{10.1}$$

$$P_i^{(n)}(0) = 1 \quad \forall\, i \text{ and } n \tag{10.2}$$

$$\lim_{T\to\infty} P_i^{(n)}(T) = 0 \quad \forall\, i \text{ and } n \tag{10.3}$$

10.2. Binomial lattice specification

HL make use of a discrete time framework within which the stochastic process describing the evolution of the discount function is represented by a binomial lattice. At time n, $P_i^{(n)}(\cdot)$ represents the discount function after i upstate moves and $(n-i)$ downstate moves. During the $(n+1)^{th}$ period, that is from time n to time $(n+1)$, the discount function may again be subject to one of two moves, an upstate move or a downstate move. Therefore, given the time n discount function $P_i^{(n)}(\cdot)$, two possible discount functions may occur at time $(n+1)$, $P_{i+1}^{(n+1)}(\cdot)$ or $P_i^{(n+1)}(\cdot)$ (see Figure 10.1). Within such a framework we have:

- At each time n, there are $(n+1)$ possible states, denoted by i, $i = 0, \ldots, n$.
- The discount function in each state is independent of the path followed to get there. It is defined by the number of upstate and downstate moves only.

The price of each discount bond follows a binomial process where the step size is time-dependent. This feature greatly increases the explanatory power of the binomial lattice approach. When modelling the term structure with a view to pricing interest rate contingent claims, we are concerned with the movement of interest rates of various maturities relative to each other – that

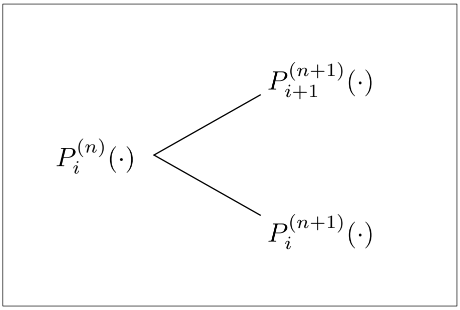

FIGURE 10.1. Binomial tree showing possible discount bond prices after one time step.

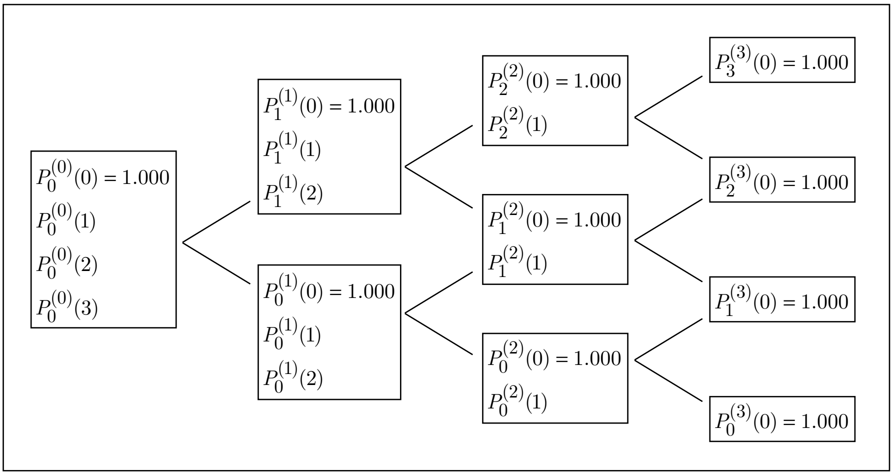

FIGURE 10.2. Binomial tree is used to model the whole term structure, not only the short-term interest rate.

is the relative movement of different maturity discount bonds. For this reason the binomial lattice is used to model the entire term structure instead of just a specific bond. The time-dependent step size ensures the convergence, at maturity, of the bond price to unity.

The binomial lattice approach imparts the following characteristics on the stochastic process of the bond price:

- Bond price uncertainty is small near bond maturity and in the immediate future.
- Bond price uncertainty increases with distance from these two points.

These characteristics are a result of two factors:

- For longer time horizons, the number of variations and hence uncertainty associated with the term structure increases.
- As the time horizon approaches bond maturity, price uncertainty decreases since the bond price must converge to unity at maturity.

For a given bond, as the time horizon increases, the term structure uncertainty increases resulting in greater price variance. This is accompanied by the bond approaching maturity. At some future time horizon, the pull-to-par effect dominates, resulting in a decrease in bond price variability.

10.3. Arbitrage-free interest rate evolution

The movement of the term structure must be constrained to ensure that no arbitrage principles are not violated. This translates to constraints on the binomial lattice modelling the term structure evolution. HL impose an additional restriction which simplifies the construction of the pricing lattice. The perturbation functions and implied binomial probability ($h(T)$, $h^*(T)$ and π respectively[1]) are taken to be independent of time n and state i. This is equivalent to the continuous time constant volatility assumption. Allowing this functional dependence would lead to a more general arbitrage-free model.

10.3.1. The perturbation functions. Consider the time n, state i discount function $P_i^{(n)}(T)$. If there is no interest rate risk/uncertainty over the next time period, then the time $(n+1)$ discount functions and hence entire term structure must be the same in the upstate and downstate. To prevent arbitrage opportunities, the realised (time $(n+1)$) discount function must be the (time n) implied forward discount function, $F_i^{(n)}(T)$:

$$F_i^{(n)}(T) = P_i^{(n+1)}(T) = P_{i+1}^{(n+1)}(T) = \frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)} \quad T = 0, 1, \dots \tag{10.4}$$

In this riskless world, if the realised discount function differs from the implied forward discount function, arbitrage opportunities exist. Hence, modelling term structure uncertainty reduces to determining the perturbation of the next period discount function from the implied forward discount function.

Let $h(T)$ and $h^*(T)$ be two perturbation functions where:

$$P_{i+1}^{(n+1)}(T) = \frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)} h(T) \tag{10.5}$$

$$P_i^{(n+1)}(T) = \frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)} h^*(T) \tag{10.6}$$

define the upstate and downstate perturbations respectively. These two perturbations specify the deviation, from the implied forward discount function, of the upstate and downstate discount functions. Hence these functions give an indication of the difference between the upstate and downstate discount functions in the next period. For $h(T) >> 1$ ($h^*(T) << 1$) for all bond maturities, the bond prices will consistently rise in the upstate (fall in the downstate).

Conditions (10.1) to (10.3) imply that:

$$h(T),\ h^*(T) > 0 \quad \forall T$$

[1] These variable names are given here for completeness only; the investigation which follows defines these variables correctly in the appropriate context.

and[2]

$$h(0) = h^*(0) = 1$$

The magnitude of the perturbation depends on bond maturity and hence $h(\cdot)$ and $h^*(\cdot)$ are functions of T. To construct the binomial lattice determining the term structure movement, we require the set of perturbation functions $\{h(T),\, h^*(T) : \forall T\}$ and the initial discount function $P_0^{(0)}(T) = P(T)$.

10.3.2. The implied binomial probability. As stated above, knowing the perturbation functions and initial discount function, allows us to construct the term structure movement. However, the modelled evolution of various maturity interest rates relative to each other must be such that profitable arbitrage opportunities do not arise. Hence consider a portfolio of two different maturity discount bonds held in proportions such that risk-free rate of return is realised over the next time period. To preclude profitable arbitrage, this risk-free rate of return must equal the return on a one period discount bond. This requirement implies a restriction on the perturbation functions at each node (n, i) of the lattice. These restrictions can be determined by constructing the above mentioned risk-free portfolio:

At any time n, state i we construct a portfolio with:

- 1 discount bond with maturity T,
- ξ discount bonds with maturity t.

Suppressing the notational dependence on time and state (n, i), the value of the portfolio is:

$$V = P(T) + \xi P(t)$$

Depending on whether an upstate or downstate move occurs during the next time period, use (10.5) or (10.6) to revalue the portfolio:

$$V(\text{upstate}) = \frac{P(T)h(T-1) + \xi P(t)h(t-1)}{P(1)} \tag{10.7}$$

$$V(\text{downstate}) = \frac{P(T)h^*(T-1) + \xi P(t)h^*(t-1)}{P(1)} \tag{10.8}$$

[2]Consider (10.5) and (10.6) with $T = 0$

$$\begin{aligned} P_{i+1}^{(n)}(0) &= \frac{P_i^{(n-1)}(1)}{P_i^{(n-1)}(1)}\, h(0) = 1 \quad \Rightarrow\ h(0) = 1 \\ P_i^{(n)}(0) &= \frac{P_i^{(n-1)}(1)}{P_i^{(n-1)}(1)}\, h^*(0) = 1 \quad \Rightarrow\ h^*(0) = 1 \end{aligned}$$

since all bonds mature with face value 1.

For a risk-free portfolio, we require $V(\text{upstate}) = V(\text{downstate})$, so:

$$P(T)h(T-1) + \xi P(t)h(t-1) = P(T)h^*(T-1) + \xi P(t)h^*(t-1)$$

$$\Rightarrow \xi = \frac{P(T)h^*(T-1) - P(T)h(T-1)}{P(t)h(t-1) - P(t)h^*(t-1)} \tag{10.9}$$

To prevent profitable arbitrage, a risk-free portfolio must yield, after one period, the return on a one period discount bond, i.e. $1/P(1)$. Therefore, we require[3]:

$$\left(\frac{1}{P(1)}\right)\big(P(T) + \xi P(t)\big) = \frac{P(T)h(T-1) + \xi P(t)h(t-1)}{P(1)}$$

$$P(T) + \xi P(t) = P(T)h(T-1) + \xi P(t)h(t-1) \tag{10.10}$$

Substituting the calculated value of ξ from (10.9) we have

$$\begin{aligned} &P(T)h(T-1) + \xi P(t)h(t-1) \\ &= \; P(T)h(T-1) + \frac{P(T)\;[h^*(T-1) - h(T-1)]}{P(t)\;[h(t-1) - h^*(t-1)]}\,P(t)h(t-1) \end{aligned}$$

$$= \; \frac{P(T)\;[h(t-1)h^*(T-1) - h(T-1)h^*(t-1)]}{h(t-1) - h^*(t-1)} \tag{10.11}$$

Substituting (10.9) and (10.11) into (10.10):

$$\begin{aligned} &P(T) + \frac{P(T)\;[h^*(T-1) - h(T-1)]}{P(t)\;[h(t-1) - h^*(t-1)]}\,P(t) \\ &\quad = \frac{P(T)\;[h(t-1)h^*(T-1) - h(T-1)h^*(t-1)]}{h(t-1) - h^*(t-1)} \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad & h(t-1) - h(t-1)h^*(T-1) - h^*(t-1) \\ & = h(T-1) - h(T-1)h^*(t-1) - h^*(T-1) \end{aligned}$$

---

[3]By the choice of ξ as the value resulting in equal portfolio values in the upstate and downstate, we may also use the portfolio value after an upstate move to give

$$P(T) + \xi P(t) = P(T)h^*(T-1) + \xi P(t)h^*(t-1)$$

in place of (10.10).

Adding $h^*(T-1)h^*(t-1)$ to both sides and factorising:

$$\begin{aligned} & h(t-1)\,[1-h^*(T-1)] - h^*(t-1)\,[1-h^*(T-1)] \\ & = h(T-1)\,[1-h^*(t-1)] - h^*(T-1)\,[1-h^*(t-1)] \end{aligned}$$

$$\Rightarrow \quad \frac{1-h^*(T-1)}{h(T-1)-h^*(T-1)} = \frac{1-h^*(t-1)}{h(t-1)-h^*(t-1)} \qquad \forall T,\, t>0$$

The LHS of (10.12) is a function of T only, while the RHS is a function of t only. This can only hold true if

$$\frac{1-h^*(T-1)}{h(T-1)-h^*(T-1)} = \frac{1-h^*(t-1)}{h(t-1)-h^*(t-1)} = \pi \qquad \forall T,\, t>0 \tag{10.12}$$

where π is some constant. More generally, we write:

$$\frac{1-h^*(T)}{h(T)-h^*(T)} = \pi \qquad \forall T>0$$

and so

$$\begin{aligned} 1-h^*(T) &= \pi h(T) - \pi h^*(T) \\ \Rightarrow 1 &= \pi h(T) + (1-\pi)h^*(T) \qquad \forall T \geq 0 \text{ and } n,i\ >0 \end{aligned} \tag{10.13}$$

If the condition in (10.13) is satisfied across all bond maturities, profitable arbitrage opportunities are precluded.

The constant π is independent of bond maturity T and the initial discount function $P(T)$, but it may depend on time n and state i. We refer to π as the implied binomial probability.

10.3.3. Comparison to traditional option pricing approach. Consider solving (10.5) and (10.6) for $h(T)$ and $h^*(T)$ respectively:

$$\begin{aligned} h(T) &= \frac{P_{i+1}^{(n+1)}(T)\,P_i^{(n)}(1)}{P_i^{(n)}(T+1)} \\ h^*(T) &= \frac{P_i^{(n+1)}(T)\,P_i^{(n)}(1)}{P_i^{(n)}(T+1)} \end{aligned}$$

Substituting the above values into (10.13), we have:

$$1 = \pi\frac{P_{i+1}^{(n+1)}(T)\,P_i^{(n)}(1)}{P_i^{(n)}(T+1)} + (1-\pi)\frac{P_i^{(n+1)}(T)\,P_i^{(n)}(1)}{P_i^{(n)}(T+1)}$$

$$\Rightarrow P_i^{(n)}(T+1) = \left[\pi\, P_{i+1}^{(n+1)}(T) + (1-\pi)P_i^{(n+1)}(T)\right]\, P_i^{(n)}(1) \tag{10.14}$$

which is consistent with the traditional binomial tree approach to option pricing introduced by Cox, Ross and Rubinstein [**19**]. In their approach the price at time n is the weighted sum of prices at time $(n+1)$, discounted by the risk-free rate over this one time period. The appropriate weights are the probabilities of the time $(n+1)$ prices being attained. For this reason π may be viewed as a risk-neutral probability.

Let us examine (10.14) in more detail. Rearranging terms, we have:

$$\begin{aligned}
\frac{P_i^{(n)}(T)}{P_i^{(n)}(1)} &= \pi\left[P_{i+1}^{(n+1)}(T-1) - P_i^{(n+1)}(T-1)\right] \\
&\quad + P_i^{(n+1)}(T-1) \\
\frac{1}{P_i^{(n)}(1)} - \frac{P_i^{(n+1)}(T-1)}{P_i^{(n)}(T)} &= \pi\left[\frac{P_{i+1}^{(n+1)}(T-1)}{P_i^{(n)}(T)} - \frac{P_i^{(n+1)}(T-1)}{P_i^{(n)}(T)}\right]
\end{aligned}$$

where

$$\frac{P_{i+1}^{(n+1)}(T-1)}{P_i^{(n)}(T)} = u \qquad \text{is the percentage up move,}$$

$$\frac{P_i^{(n+1)}(T-1)}{P_i^{(n)}(T)} = d \qquad \text{is the percentage down move,}$$

and

$$\frac{1}{P_i^{(n)}(1)} = r \qquad \text{is the single period risk-free return.}$$

Hence

$$\begin{aligned}
r - d &= \pi(u-d) \\
\pi &= \frac{r-d}{u-d}
\end{aligned}$$

and π may be interpreted as the extent of the downstate move with respect to the spread between an upstate and downstate move. Large π implies a general price decrease over the next time period, while for values of π close to zero, a general price rise may be expected over the next time period. The required no arbitrage condition of equation (10.12) implies that this ratio (i.e. the general increase/decrease in the bond price over the next time period) must be the same for all bond maturities[4] T.

[4]This is the discrete time equivalent of the continuous time condition that to preclude arbitrage opportunities the instantaneous return on bonds of all maturities must be the same.

Cox, Ross and Rubinstein applied a constant discount rate at all points of the binomial tree. In this model the one-period discount rate is time and state dependent. This is due to the one-period rates, $P_i^{(n)}(1)$ for all n, i being determined by the initial term structure and hence endogenised by the no arbitrage methodology.

10.3.4. Constraints to ensure path independence. As we construct the binomial tree/lattice, the discount function evolves according to the number of upstate and downstate moves only, not their sequence. This condition implies constraints on h^*, h, and π so that at any time n, state i, an upward move, followed by a downward move is equivalent to a downward move followed by an upward move. We examine the restrictions required to ensure this path independence.

First, consider a downward move, followed by an upward move to get from $P_i^{(n)}(T+2)$ to $P_{i+1}^{(n+2)}(T)$. From (10.6) the downward move is represented by:

$$P_i^{(n+1)}(T+1) = \frac{P_i^{(n)}(T+2)}{P_i^{(n)}(1)}\, h^*(T+1)$$

and then from (10.5) the upward move is:

$$P_{i+1}^{(n+2)}(T) = \frac{P_i^{(n+1)}(T+1)}{P_i^{(n+1)}(1)}\, h(T)$$

Hence we have:

$$P_{i+1}^{(n+2)}(T) = \frac{P_i^{(n)}(T+2)\, h^*(T+1)\, h(T)}{P_i^{(n)}(1)\, P_i^{(n+1)}(1)}$$

Also, we may write:

$$\begin{aligned} P_i^{(n+1)}(1) &= \frac{P_i^{(n)}(2)}{P_i^{(n)}(1)}\, h^*(1) \\ \Rightarrow P_{i+1}^{(n+2)}(T) &= \frac{P_i^{(n)}(T+2)\, h^*(T+1)\, h(T)}{P_i^{(n)}(2)\, h^*(1)} \qquad (10.15) \end{aligned}$$

Similarly, an upward move, followed by a downward move yields:

$$\begin{aligned} P_{i+1}^{(n+1)}(T+1) &= \frac{P_i^{(n)}(T+2)}{P_i^{(n)}(1)}\, h(T+1) \\ P_{i+1}^{(n+2)}(T) &= \frac{P_{i+1}^{(n+1)}(T+1)}{P_{i+1}^{(n+1)}(1)}\, h^*(T) \end{aligned}$$

$$\begin{aligned}
\text{and} \quad P_{i+1}^{(n+1)}(1) &= \frac{P_i^{(n)}(2)}{P_i^{(n)}(1)}\, h(1) \\
\Rightarrow P_{i+1}^{(n+2)}(T) &= \frac{P_i^{(n)}(T+2)\, h(T+1)\, h^*(T)}{P_i^{(n)}(2)\, h(1)} \qquad (10.16)
\end{aligned}$$

The path independence condition means that (10.15) must equal (10.16) so:

$$\begin{aligned}
\frac{P_i^{(n)}(T+2)\, h^*(T+1)\, h(T)}{P_i^{(n)}(2)\, h^*(1)} &= \frac{P_i^{(n)}(T+2)\, h(T+1)\, h^*(T)}{P_i^{(n)}(2)\, h(1)} \\
\Rightarrow h^*(T+1)\, h(T)\, h(1) &= h(T+1)\, h^*(T)\, h^*(1) \qquad (10.17)
\end{aligned}$$

From (10.13) we may express $h^*(\cdot)$ as a function of $h(\cdot)$:

$$h^*(T) = \frac{1-\pi\, h(T)}{1-\pi} \qquad \forall\, T$$

so (10.17) simplifies as:

$$\begin{aligned}
\frac{1-\pi\, h(T+1)}{1-\pi}\, h(T)\, h(1) &= h(T+1)\left(\frac{1-\pi\, h(T)}{1-\pi}\right)\left(\frac{1-\pi\, h(1)}{1-\pi}\right) \\
\frac{1-\pi\, h(T+1)}{h(T+1)} &= \left(\frac{1-\pi\, h(T)}{h(T)}\right)\left(\frac{1-\pi\, h(1)}{(1-\pi)h(1)}\right) \\
\frac{1}{h(T+1)} - \pi &= \left(\frac{1}{h(T)} - \pi\right)\left(\frac{1-\pi\, h(1)}{(1-\pi)h(1)}\right) \\
\frac{1}{h(T+1)} &= \frac{1-\pi\, h(1)}{(1-\pi)\, h(1)\, h(T)} - \frac{\pi(1-\pi\, h(1))}{(1-\pi)h(1)} + \pi \\
&= \frac{1-\pi\, h(1)}{(1-\pi)\, h(1)\, h(T)} + \frac{\pi(h(1)-1)}{(1-\pi)h(1)}
\end{aligned}$$

Letting

$$\delta = \frac{1-\pi\, h(1)}{(1-\pi)\, h(1)} \quad \text{and} \quad \gamma = \frac{\pi(h(1)-1)}{(1-\pi)h(1)} \qquad (10.18)$$

we may represent the condition guaranteeing path independence as a first-order linear difference equation of the form:

$$\frac{1}{h(T+1)} = \frac{\delta}{h(T)} + \gamma \qquad (10.19)$$

10.3.5. Solution to the first-order linear difference equation. Let $g_r = \frac{1}{h(T)}$ and so $g_{r+1} = \frac{1}{h(T+1)}$ and (10.19) may be written as:

$$g_{r+1} = g_r \delta + \gamma \qquad (10.20)$$

First, let $x_{r+1} = E(x_r) = x_r + \Delta x$ and consider the solution to the homogeneous part of the difference equation:

$$(E - \delta)g_r = 0$$

which has the root δ and the solution[5] is $g_r^H = k_1\delta^r$, where k_1 is a constant.

Now calculate the particular part of the solution:

$$\begin{aligned}(E - \delta)g_r &= \gamma \\ (E - \delta)(E - 1)g_r &= (E - 1)\gamma \\ &= \gamma - \gamma \;=\; 0\end{aligned}$$

This equation has roots δ and 1, so the particular solution[6] is $g_r^P = k_2\delta^r + k_3$ where k_2 and k_3 are again constants. Since δ^r is included in the homogeneous part of the solution, we exclude it and substitute $g_r^P = k_3$ into (10.20) to solve for the constant. Hence[7]:

$$\begin{aligned}k_3 - k_3\delta &= \gamma \\ \Rightarrow k_3 &= \frac{\gamma}{1-\delta} \;=\; \pi\end{aligned}$$

So the solution to (10.19) is

$$\begin{aligned}g_r &= g_r^P + g_r^H \\ &= \pi + k_1\delta^r \\ \Rightarrow h(T) &= \frac{1}{\pi + k_1\delta^T} \qquad (10.21)\end{aligned}$$

The initial condition on h requires $h(0) = 1$, so from (10.21) $k_1 = 1 - \pi$ and the unique solution is determined as:

$$h(T) = \frac{1}{\pi + (1-\pi)\,\delta^T} \quad \text{for } T \geq 0 \qquad (10.22)$$

Equation (10.13) gives the relationship between $h(T)$ and $h^*(T)$ required to preclude profitable arbitrage. It may be used in conjunction with (10.22) to solve for an $h^*(T)$ which maintains the no arbitrage equilibrium. Substituting (10.22) into (10.13) to have:

---

[5]Here the superscript H denotes the homogeneous part of the solution.

[6]Here the superscript P denotes the particular part of the solution.

[7]Making use of the definitions of γ and δ in (10.18) we calculate $\gamma/(1-\delta) = \pi$

$$
\begin{aligned}
1 &= \frac{\pi}{\pi + (1-\pi)\,\delta^T} + (1-\pi)h^*(T) \\
\Rightarrow h^*(T) &= \frac{1}{1-\pi} - \frac{\pi}{(1-\pi)(\pi + (1-\pi)\delta^T)} \\
&= \frac{\delta^T}{\pi + (1-\pi)\delta^T} \qquad (10.23)
\end{aligned}
$$

For a given pair of constraints, π and δ (which corresponds to a specification of $h(1)$), the unique arbitrage-free model is fully specified by (10.5), (10.6), (10.22) and (10.23). Here π is the implied binomial probability and δ determines the spread between the perturbation functions $h(\cdot)$ and $h^*(\cdot)$.

10.4. Relationship to Vasicek and CIR models

The Vasicek [**50**] and CIR [**18**] models describe the instantaneous short-term interest rate by means of a stochastic process. Within the discrete time HL model, the short-term interest rate equivalent is the one period rate. To make a meaningful comparison of these models we need to find the stochastic process described by the evolution of this rate in the binomial lattice.

10.4.1. Calculating the continuous time equivalent.

10.4.1.1. *General functional form.* Rebonato [**45**] presents a simple analysis by which the continuous time equivalent of a discrete time model, modelled within a binomial lattice, may be found.

At each time step t in a binomial lattice, there are $t+1$ possible states of the world and hence $t+1$ possible values of the one period rate. Consider time $t = 1$, there are two possible states of the world and interest rates, denoted $r(1,1)$ and $r(1,-1)$. Given the assumption that the short-term interest rate follows a Gaussian process, the standard deviation of the time $t = 1$ one-period rate may be represented as:

$$\sigma\sqrt{\Delta t} = \frac{r(1,1) - r(1,-1)}{2} \quad \Rightarrow \quad r(1,1) = r(1,-1) + 2\sigma\sqrt{\Delta t} \qquad (10.24)$$

where σ is the absolute volatility of the one period rate.

Let $r_m(1)$ be the median interest rate at time $t = 1$; hence:

$$
\begin{aligned}
r_m(1) &= \tfrac{1}{2}\left[r(1,1) + r(1,-1)\right] \\
\Rightarrow \quad r(1,1) &= r_m(1) + \sigma\sqrt{\Delta t} \\
\text{and} \quad r(1,-1) &= r_m(1) - \sigma\sqrt{\Delta t}
\end{aligned}
$$

and so, in continuous time, we may write:

$$r(t) = u(t) + \sigma z(t)$$

where $u(t)$ is time t median of the short-term interest rate distribution, σ is the constant short-term interest rate volatility and $z(t)$ is a standard Brownian motion. Now, apply Ito's Lemma to determine the stochastic process for the short-term interest rate, $r(t) = r(t, z(t))$:

$$\begin{aligned} dr &= \frac{\partial r}{\partial t}\,dt + \frac{\partial r}{\partial z}\,dz + \tfrac{1}{2}\frac{\partial^2 r}{\partial z^2}\,dzdz \\ &= \frac{\partial u(t)}{\partial t}\,dt + \sigma dz \end{aligned}$$

Letting $\theta(t) = \frac{\partial u(t)}{\partial t}$, the process for the short-term interest rate may be expressed as:

$$dr = \theta(t)\,dt + \sigma dz \tag{10.25}$$

where $\theta(t)$ is a function of the initial term structure. This is the case since $u(t)$, the median of the short-term interest rate, is determined as part of the binomial tree calibration process.

10.4.1.2. *Specific functional form of $\theta(t)$ dependent on lattice parameters.* Now we examine how this drift function, $\theta(t)$, is dependent on the specific parameters of the binomial tree. First, we make use of (10.5) and (10.6) to determine the discount function at time n, state i, in terms of the initial discount function and the perturbations required to get there. Equations (10.5) and (10.6) are:

$$P_{i+1}^{(n+1)}(T) = \frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)}\,h(T) \tag{10.26}$$

$$P_{i}^{(n+1)}(T) = \frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)}\,h^*(T) \tag{10.27}$$

Repeated application of (10.27) yields:

$$\begin{aligned} P_i^{(n)}(T) &= \frac{P_i^{(n-1)}(T+1)}{P_i^{(n-1)}(1)}\,h^*(T) \\ &= \frac{P^{(n-2)}(T+2)\,h^*(T+1)\,h^*(T)}{P_i^{(n-2)}(1)\,P_i^{(n-1)}(1)} \\ &= \frac{P^{(n-3)}(T+3)\,h^*(T+2)\,h^*(T+1)\,h^*(T)}{P_i^{(n-3)}(1)\,P_i^{(n-2)}(1)\,P_i^{(n-1)}(1)} \end{aligned}$$

and

$$P_i^{(n-1)}(1) = \frac{P_i^{(n-2)}(2)}{P_i^{(n-2)}(1)}\, h^*(1) \qquad\qquad P_i^{(n-2)}(2) = \frac{P_i^{(n-3)}(3)}{P_i^{(n-3)}(1)}\, h^*(2)$$

Hence:

$$P_i^{(n)}(T) = \frac{P_i^{(n-3)}(T+3)\, h^*(T+2)\, h^*(T+1)\, h^*(T)}{P_i^{(n-3)}(3)\, h^*(2)\, h^*(1)}$$

Eventually we arrive at the formula:

$$\begin{aligned} P_i^{(n)}(T) &= \frac{P(T+n)}{P(n)} \\ &\times \frac{h^*(T+n-1)\, h^*(T+n-2)\ldots h^*(T+i)\, h(T+i-1)\ldots h(T)}{h^*(n-1)\, h^*(n-2)\ldots h^*(i)\, h(i-1)\ldots h(1)} \end{aligned} \tag{10.28}$$

The above combination of perturbation functions results from the i upstate moves and $(n-i)$ downstate moves required to reach the i^{th} state at time n. The order of the moves is not important since the time n, state i, discount function is path-independent. Equation (10.28) can be used without loss of generality.

From (10.22) and (10.23) we have:

$$h^*(T) = \frac{\delta^T}{\pi + (1-\pi)\,\delta^T} = \delta^T\, h(T)$$

Hence (10.28) may be written as:

$$\begin{aligned} P_i^{(n)}(T) &= \frac{P(T+n)}{P(n)} \\ &\times \frac{h(T+n-1)\, h(T+n-2)\ldots h(T+i)\, h(T+i-1)\ldots}{h(n-1)\, h(n-2)\ldots h(i)\, h(i-1)\ldots h(1)\, \delta^{n-1}\, \delta^{n-2}\ldots \delta^i} \\ &\times \frac{\ldots h(T)\, \delta^{T+n-1}\, \delta^{T+n-2}\ldots \delta^{T+i}}{h(n-1)\, h(n-2)\ldots h(i)\, h(i-1)\ldots h(1)\, \delta^{n-1}\, \delta^{n-2}\ldots \delta^i} \\ &=^8 \frac{P(T+n)}{P(n)}\, \frac{h(T+n-1)\, h(T+n-2)\ldots h(T)}{h(n-1)\, h(n-2)\ldots h(1)}\, \delta^{T(n-i)} \end{aligned} \tag{10.29}$$

where the exponent of δ is consistent with requiring $(n-i)$ downstate moves (and i upstate moves) to reach time n, state i.

Now, consider the special case of a one-period bond i.e. $T = 1$. From (10.29) we have:

[8]Considering the δ terms in the numerator, we have:

$$\delta^{T+n-1}\, \delta^{T+n-2}\ldots \delta^{T+i} = \delta^{T+n-1+T+n-2+\ldots+T+i}$$

Here the exponent may be written as:

$$\begin{aligned} P_i^{(n)}(1) &= \frac{P(n+1)}{P(n)} \frac{h(n)\,h(n-1)\ldots h(1)}{h(n-1)\,h(n-2)\ldots h(1)}\,\delta^{n-i} \\ &= \frac{P(n+1)}{P(n)}\,h(n)\,\delta^{n-i} \end{aligned}$$

But, from (10.22) we have:

$$h(T) = \frac{1}{\pi + (1-\pi)\,\delta^T}$$

Hence:

$$P_i^{(n)}(1) = \frac{P(n+1)}{P(n)}\,\frac{\delta^{n-i}}{\pi + (1-\pi)\,\delta^n}$$

To represent the interest rate term structure in terms of yields as opposed to discount functions, let $r(T)$ be the continuously compounded yield on a discount bond of maturity T; then:

$$\begin{aligned} P(T) &= e^{-r(T)T} \\ \Rightarrow r(T) &= -\frac{\ln P(T)}{T} \end{aligned} \tag{10.30}$$

$$\begin{aligned} T+n-1+T+n-2+\ldots+T+i &= \sum_{k=T+i}^{T+n-1} k \\ &= \sum_{k=1}^{T+n-1} k - \sum_{k=1}^{T+i-1} k \\ &= \tfrac{1}{2}\left[(T+n)(T+n-1)-(T+i)(T+i-1)\right] \end{aligned}$$

Similarly, the exponent of the δ terms in the denominator may be expressed as:

$$\begin{aligned} n-1+n-2+\ldots+i &= \sum_{k=i}^{n-1} k \\ &= \tfrac{1}{2}\left[n(n-1)-i(i-1)\right] \end{aligned}$$

Then, to simplify, we subtract the exponent of δ in the denominator from the exponent of δ in the numerator to yield:

$$\tfrac{1}{2}\left[(T+n)(T+n-1)-(T+i)(T+i-1)\right] - \tfrac{1}{2}\left[n(n-1)-i(i-1)\right] = T(n-i)$$

and so the time n, state i, one period yield may be expressed as:

$$
\begin{aligned}
r_i^{(n)}(1) &= -\ln P_i^{(n)}(1) \\
&= -\ln \left[\frac{P(n+1)}{P(n)} \, \frac{\delta^{n-i}}{\pi + (1-\pi)\,\delta^n} \right] \\
&= \ln \left[\frac{P(n)}{P(n+1)} \right] + \ln \left[\frac{\pi + (1-\pi)\,\delta^n}{\delta^{n-i}} \right] \\
&= \ln \left[\frac{P(n)}{P(n+1)} \right] + \ln \left[\pi\,\delta^{-n+i} + (1-\pi)\,\delta^i \right] \\
&= \ln \left[\frac{P(n)}{P(n+1)} \right] + \ln \left[\pi\,\delta^{-n} + (1-\pi) \right] + i \ln \delta \qquad (10.31)
\end{aligned}
$$

The value of i, that is the possible state at time n, has a Binomial distribution[9] with probability q. Hence at time n the mean value of i is $\mu(i) = nq$. Since $r_i^{(n)}(1)$ is unique for every state and time, it also has a Binomial distribution in i for each time n. The mean of this distribution, that is the mean short-term interest rate at time n, may be calculated from (10.31) as:

$$
\mu\left(r_i^{(n)}(1)\right) = \ln \left[\frac{P(n)}{P(n+1)} \right] + \ln \left[\pi\,\delta^{-n} + (1-\pi) \right] + n\,q \ln \delta \qquad (10.32)
$$

The first term in (10.32) is the implied forward rate, while the other two terms give the bias introduced by uncertainty, i.e. δ which gives the ratio of the upstate and downstate perturbations. Consider the drift of the short-term interest rate $\theta(t)$, in (10.25). This drift is chosen so as to fit the initial observed term structure. This dependence on the initial term structure is represented by the first term in (10.32), i.e. the implied forward rate. The time dependence of $\theta(t)$ is indicated by the parameter n, in each of the three terms of (10.32).

The variance of the short-term interest rate may be calculated from (10.31) as follows:

[9] At each step of the Binomial tree the movement of the discount function is subject to a Bernoulli trial. A Bernoulli trial has one of two possible outcomes – success (1) or failure (0), where q denotes the probability of success. The Binomial distribution is made up of n identical Bernoulli trials. It has the following characteristics:

$$
\text{mean } = n\,q \qquad\qquad \text{variance } = n\,q(1-q)
$$

$$
\begin{aligned}
\operatorname{var}\left(r_i^{(n)}(1)\right) &= \mathbb{E}\left[\left(r_i^{(n)}(1) - \mathbb{E}\left[r_i^{(n)}(1)\right]\right)^2\right] \\
&= \mathbb{E}\left[(i \ln\delta - n\,q\ln\delta)^2\right] \\
&= (\ln\delta)^2\ \mathbb{E}\left[(i - n\,q)^2\right] \\
&= (\ln\delta)^2\ \operatorname{var}(i) \\
&= n\,q(1-q)(\ln\delta)^2 \qquad (10.33)
\end{aligned}
$$

As the spread between the up and down perturbation functions ($h(\cdot)$ and $h^*(\cdot)$) increases, so the value of δ decreases[10]. An increase in this spread implies an increase in variance, hence the variance is negatively related to δ.

Consider the volatility σ, of the short-term interest rate defined in (10.25). This is the instantaneous volatility, hence for a specified time period the volatility is $\sigma\sqrt{t}$. From (10.33) the corresponding time n short-term interest rate volatility is $\ln\delta\sqrt{q(1-q)}\sqrt{n}$. Therefore the constant value σ is represented by $\ln\delta\sqrt{q(1-q)}$ since δ and q are constant parameters[11] associated with the binomial lattice.

10.4.2. Comparing the modelling techniques. By calculating the mean and variance of the short-term interest rate, we have specified the stochastic process that governs its evolution. This stochastic process depends on information contained in the initial term structure; therefore the future evolution of the short-term interest rate is determined by the initial term structure. Traditional one factor models, such as the Vasicek and CIR[12] models, do not fit the initial term structure exactly since they specify the short-term interest

[10] δ is defined as $h^*(\cdot)/h(\cdot)$ that is, the down perturbation function divided by the up perturbation function, hence $\delta \leq 1$ for all values of $h(\cdot)$ and $h^*(\cdot)$. As the spread between the two functions increases, so δ tends to zero. Therefore:

$$
\begin{aligned}
& \ln\delta \leq 0 \qquad \forall\ \delta \\
\text{and} \quad & \lim_{\delta\to 0}\ln\delta = -\infty \\
\text{so} \quad & \lim_{\delta\to 0}(\ln\delta)^2 = \infty
\end{aligned}
$$

Hence the increasing value of the variance with decreasing value of δ.

[11] Parameter q is the binomial probability corresponding to the risk-neutral probability π characterising the lattice.

[12] In the most general form, the process used by the traditional Vasicek and CIR models is:

$$dr = \kappa(\theta - r)\,dt + \sigma r^\beta dz$$

where κ, θ and σ are constants. In the Vasicek model $\beta = 0$ while in the CIR model $\beta = \frac{1}{2}$. Since none of the parameters are time-dependent, only the initial short-term interest rate can be fitted to a rate observed in the market.

rate process exogenously. These two different approaches (that is the Vasicek/CIR on one hand and HL on the other) may be justified since they are each applied for a different purpose. The traditional one-factor models, such as the Vasicek and CIR models, endogenise the equilibrium term structure by attempting to determine a short-term interest rate process that generates a meaningful equilibrium term structure. The purpose of the approach taken by HL is not to determine an equilibrium term structure, but rather to price contingent claims consistently with respect to the initial term structure. For this reason the short-term interest rate movements incorporate information from the initial term structure.

10.5. Pricing contingent claims

The binomial lattice used to price contingent claims is characterised as follows:

C	–	interest rate contingent claim,
$C(n,i)$	–	unique price of the contingent claim defined at each node(n,i),
T	–	expiry time of the contingent claim,
$\{f(i)\}$	–	set of payoffs of the contingent claim at expiry $0 \leq i \leq T$ (since at each time$n$ there are $(n+1)$ states, hence at time T there are $(T+1)$ possible payoff values). Therefore the terminal condition is $C(T,i) = f(i),\ 0 \leq i \leq T$,
$L(n,i),\ U(n,i)$	–	lower and upper bounds on the contingent claim price such that $L(n,i) \leq C(n,i) \leq U(n,i)$,
$X(n,i)$	–	amount paid by the contingent claim at time n, state i, $1 \leq n \leq T$.

10.5.1. No arbitrage contingent claim price. Let V be the value of a risk-free portfolio made up of one discount bond with maturity T and ξ of asset C. Consider the time n, state i value of the portfolio:

$$V = P_i^{(n)}(T) + \xi\, C(n,i) \tag{10.34}$$

This portfolio is subject to an upstate move. Hence the price of the discount bond becomes:

$$P_{i+1}^{(n+1)}(T-1) = \frac{P_i^{(n)}(T)}{P_i^{(n)}(1)}\, h(T-1)$$

and the value of the portfolio may be expressed as:

$$\begin{aligned} V(\text{upstate}) &= P_{i+1}^{(n+1)}(T-1) + \xi\, C(n+1,i+1) \\ &= \frac{P_i^{(n)}(T)}{P_i^{(n)}(1)}\, h(T-1) + \xi\, C(n+1,i+1) \end{aligned}$$

Similarly, if the portfolio is subject to a downward move its value is:

$$\begin{aligned} V(\text{downstate}) &= P_{i}^{(n+1)}(T-1) + \xi\, C(n+1,i) \\ &= \frac{P_i^{(n)}(T)}{P_i^{(n)}(1)}\, h^*(T-1) + \xi\, C(n+1,i) \end{aligned}$$

By definition the portfolio is risk-free and so $V(\text{upstate}) = V(\text{downstate})$. Therefore, the amount of asset C required to ensure a risk-free portfolio is calculated as:

$$\frac{P_i^{(n)}(T)}{P_i^{(n)}(1)}\, h^*(T-1) - \frac{P_i^{(n)}(T)}{P_i^{(n)}(1)}\, h(T-1) = \xi\,[C(n+1,i+1) - C(n+1,i)]$$

$$\Rightarrow \xi = \frac{P_i^{(n)}(T)\,[h^*(T-1) - h(T-1)]}{P_i^{(n)}(1)\,[C(n+1,i+1) - C(n+1,i)]} \tag{10.35}$$

Given this value of ξ the portfolio is risk-free over one period and hence must earn the risk-free rate of return, $1/P_i^{(n)}(1)$:

$$\begin{gathered} V = V(\text{downstate})\, P_i^{(n)}(1) = V(\text{upstate})\, P_i^{(n)}(1) \\ \Rightarrow P_i^{(n)}(T) + \xi\, C(n,i) = P_i^{(n)}(T)\, h^*(T-1) + \xi\, P_i^{(n)}(1)\, C(n+1,i) \end{gathered}$$

and making use of (10.35), (10.22) and (10.23) we may solve for the time n, state i price of the contingent claim as a function of the possible contingent claim prices at time $(n+1)$:

$$\begin{aligned} C(n,i) &= P_i^{(n)}(1)\, C(n+1,i) + \frac{P_i^{(n)}(T)\,[h^*(T-1) - 1]}{\xi} \\ &= P_i^{(n)}(1)\, C(n+1,i) + \frac{P_i^{(n)}(T)\left[\frac{\delta^T}{\pi+(1-\pi)\delta^T} - 1\right] P_i^{(n)}(1)}{P_i^{(n)}(T)\left[\frac{\delta^T - 1}{\pi+(1-\pi)\delta^T}\right]} \\ &\quad \times [C(n+1,i+1) - C(n+1,i)] \\ &= P_i^{(n)}(1)\, C(n+1,i) + \pi\,[C(n+1,i+1) - C(n+1,i)]\, P_i^{(n)}(1) \\ &= [\pi\, C(n+1,i+1) + (1-\pi)\, C(n+1,i)]\, P_i^{(n)} \end{aligned} \tag{10.36}$$

Now the contingent claim may be priced by backward induction. We know the value of the contingent claim in all states i at expiry time T. The price of the contingent claim in each state i one period prior to expiry i.e. $C^*(T-1,i)$ is determined from equation (10.36). This calculated price is subject to the pre-specified boundary conditions; hence:

$$C(T-1,i) = \max\left[L(T-1,i), \min\left[C^*(T-1,i), U(T-1,i)\right]\right]$$

Iteration of this methodology produces the initial price of the contingent claim. Knowledge of the implied binomial probability and the one period discount bond price $P_i^{(n)}(1)$, at each time n, state i, allows us to price any interest rate contingent claim. We conclude that the pricing depends on the stochastic evolution of the short-term interest rate and so this model may also be referred to as a one factor model.

The parameters π and δ are not directly observable in the term structure, but are reflected in the valuation of contingent claims. An estimation procedure may be used to ensure a best fit of the model-generated (theoretical) contingent claim prices to actual observed prices. Since π and δ are characteristics of the term structure and not of the specific contingent claim, the estimated parameter values may be used to price any contingent claim.

10.6. Conclusion

This model is the first to allow a direct matching of the initial observed term structure. The model contrasts to previously developed models by considering the stochastic development of the whole term structure. Hence, each node in the binomial tree has a series of discount bond prices (or equivalently, rates of interest) of various maturities associated with it. This contrasts to, say, the BDT model (Chapter 8) where at each node, one only considers the value of the short-term interest rate applicable over the next time step.

Models such as the Vasicek [**50**] and CIR [**18**] models hypothesise a functional form for the stochastic process governing the evolution of the short-term interest rate. They then attempt to determine parameter values so as to match, as closely as possible, the market-observed term structure. HL use a different approach by using the market-observed term structure to specify the stochastic process of the short-term interest rate. This allows all securities to be priced relative to the observed term structure.

CHAPTER 11

The Heath, Jarrow and Morton Model

Heath, Jarrow and Morton (HJM) [**25**] present a unifying framework for term structure models. This framework introduces a formal elegance and generality to the interest rates modelling problem. It shows that the absence of arbitrage results in a link between the volatility of discount bonds and the drift of forward rates. In fact, in the risk-neutral world, the forward rate drift is completely determined by the specification of the discount bond volatility function. Previously developed models can be shown to be special cases of this general framework.

HJM specify an initial forward rate curve and a stochastic process describing its subsequent evolution. To ensure that the stochastic process is consistent with an arbitrage-free (and hence equilibrium) economy, it is chosen such that there exists an equivalent martingale probability measure.

The HJM methodology encompasses several new concepts:

(1) A stochastic structure is imposed on the evolution of the forward rate curve.
(2) Contingent claim prices are not dependent on the market prices of risk. This implies that inversion of the term structure to solve for these market prices of risk is not required.
(3) Evolution of the term structure is determined by the short-term interest rate, which follows a process influenced by a number of stochastic variables.

The derivation of the HJM model is rather technical in nature. It consists of a series of conditions to determine a restriction on the drift of the forward rate which ensures a risk-neutral and arbitrage-free pricing framework. In §11.2 we impose conditions to ensure well-behaved forward rate, money market and bond price processes. The relative or discounted bond price process is also defined. In §11.3 we examine necessary and sufficient conditions required to ensure the existence of a unique equivalent probability measure under which the discounted bond prices are martingales. This is equivalent to ensuring an arbitrage-free pricing framework. In §11.4 we specify a final condition ensuring a unique martingale measure across all bond maturities. It is here that the forward rate drift restriction is explicitly specified. In later sections I examine contingent claim pricing within the HJM framework, compare earlier term structure models to the HJM framework and examine conditions under which this framework gives rise to a Markovian short-term interest rate process.

11.1. Initial specifications

HJM develop their model within a continuous trading economy, with trading interval $[0, \tau]$, $\tau > 0$ fixed. Uncertainty within the economy is represented by the probability space (Ω, F, Q), where Ω represents the state space, F the σ-algebra representing all measurable events and Q the probability measure. Information becomes available over the trading period according to the filtration $\{F_t : t \in [0, \tau]\}$ which is generated by n independent Brownian motions $\{z_1(t), \ldots, z_n(t) : t \in [0, \tau]\}$ with $n \geq 1$.

Assume there exist default-free zero coupon bonds with maturities on each trading day T, $T \in [0, \tau]$. If $P(t, T)$ represents the time t price of a T-maturity bond, where $T \in [0, \tau]$ and $t \in [0, T]$, then the following must be true:

$$\begin{aligned} P(T,T) &= 1 && \forall\, T \in [0,\tau] \\ P(t,T) &> 0 && \forall\, T \in [0,\tau],\ t \in [0,T] \\ \frac{\partial \ln P(t,T)}{\partial T} &\ \text{exists} && \forall\, T \in [0,\tau],\ t \in [0,T] \end{aligned}$$

Define the time t instantaneous forward rate for time T, $T > t$ as:

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T} \quad \forall\, T \in [0,\tau],\ t \in [0,T] \tag{11.1}$$

Solving this differential equation for the bond price yields:

$$P(t,T) = \exp\Big(-\int_t^T f(t,y)\,dy\Big) \quad \forall\, T \in [0,\tau],\ t \in [0,T] \tag{11.2}$$

The short-term interest rate at time t is the instantaneous forward rate for time t, hence:

$$r(t) = f(t,t) \quad \forall\, t \in [0,\tau] \tag{11.3}$$

Alternatively, expressed in terms of the bond price[1]:

$$\begin{aligned} f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T} &= -\lim_{h \to 0}\left[\frac{\ln P(t,T+h) - \ln P(t,T)}{h}\right] \\ &= \lim_{h \to 0}\left[\frac{1}{h}\ln\frac{P(t,T)}{P(t,T+h)}\right] \end{aligned}$$

$$
\begin{aligned}
&= \lim_{h\to 0}\left[\frac{1}{h}\left(\frac{P(t,T)}{P(t,T+h)}-1\right)\right] \\
\Rightarrow r(t) &= \lim_{h\to 0}\left[\frac{1}{h}\left(\frac{1}{P(t,t+h)}-1\right)\right] \\
&= \lim_{h\to 0}\left[\frac{1-P(t,t+h)}{hP(t,t+h)}\right]
\end{aligned}
$$

Hence the short-term interest rate may be interpreted as the rate of return on an instantaneously maturing bond.

11.2. Specifications of the various processes

We present a family of stochastic processes describing the evolution of forward rates and hence uniquely determining the short-term interest rate and bond price processes. A series of conditions is presented ensuring the processes are bounded and well behaved.

11.2.1. Forward and short-term interest rate processes. Technical conditions are applied to the processes defining the short and forward interest rates as well as the money market account.

[1] Here, make use of the Taylor series expansion of the natural logarithm of a number:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad \text{where} -1 < x \leq 1$$

Consider:

$$1 + x = \frac{P(t,T)}{P(t,T+h)} \qquad \Rightarrow x = \frac{P(t,T)}{P(t,T+h)} - 1 > 0$$

By definition h is small, so $P(t,T)$ is only slightly greater than $P(t,T+h)$, and $\frac{P(t,T)}{P(t,T+h)}$ is only slightly larger than 1; hence:

$$0 < \frac{P(t,T)}{P(t,T+h)} - 1 < 1$$

Therefore applying this expansion:

$$
\begin{aligned}
\ln\frac{P(t,T)}{P(t,T+h)} &= \frac{P(t,T)}{P(t,T+h)} - 1 - \frac{1}{2}\left(\frac{P(t,T)}{P(t,T+h)} - 1\right)^2 + \dots \\
&\approx \frac{P(t,T)}{P(t,T+h)} - 1
\end{aligned}
$$

since the higher order terms are negligibly small by the definition of h.

CONDITION 1. *A family of forward rate processes.* Define a family of forward rate processes $f(t,T)$, for fixed $T \in [0,\tau]$:

$$f(t,T) - f(0,T) = \int_0^t \alpha(\omega, v, T)\, dv + \sum_{i=1}^{n} \int_0^t \sigma_i(\omega, v, T)\, dz_i(v) \qquad \forall\ 0 \le t \le T \tag{11.4}$$

where[2]:

- $\{f(0,T) : T \in [0,\tau]\}$ is a fixed, non-random initial forward rate curve, measurable as a mapping $f(0,\cdot) : ([0,\tau], \mathcal{B}[0,\tau]) \to (R, \mathcal{B})$ where $\mathcal{B}[0,\tau]$ is a Borel σ-algebra restricted to $[0,\tau]$.
- $\alpha : \Omega \times \{(t,s) : 0 \le t \le s \le T\} \to R$ is a family of drift functions jointly measurable from $F \times \mathcal{B}\{(t,s) : 0 \le t \le s \le T\} \to \mathcal{B}$, adapted and having
$$\int_0^T |\alpha(\omega, t, T)|\, dt < +\infty \qquad \text{a.e. } Q$$
- $\sigma_i : \Omega \times \{(t,s) : 0 \le t \le s \le T\} \to R$ are volatilities, jointly measurable from
$F \times \mathcal{B}\{(t,s) : 0 \le t \le s \le T\} \to \mathcal{B}$, adapted and with
$$\int_0^T \sigma_i^2(\omega, t, T)\, dt < +\infty \qquad \text{a.e. } Q \qquad \text{for } i = 1, \dots, n$$

Starting from the initial fixed forward rate curve $\{f(0,T) : T \in [0,\tau]\}$, the n independent Brownian motions determine the stochastic evolution of the whole forward curve through time. The sensitivity of the change in a given maturity forward rate to each Brownian motion, is specified by the volatility coefficients. The only restrictions imposed on the forward rate process that have economic implications are:

- time is continuous and
- stochastic movement is specified by a finite number of random shocks.

[2]Here, and in subsequent formulae ω denotes the possible dependence on the history of the Brownian motions.

CONDITION 2. *Regularity of the money market account.* Given the forward rate process in (11.4), the dynamics of the short-term interest rate may be expressed as:

$$r(t) = f(0,t) + \int_0^t \alpha(\omega, v, t)\, dv + \sum_{i=1}^{n} \int_0^t \sigma_i(\omega, v, t)\, dz_i(v) \qquad \forall\ t \in [0,\tau] \tag{11.5}$$

Now, define an accumulation factor or money market account $B(t)$, as:

$$B(t) = \exp\left(\int_0^t r(y)\, dy\right) \qquad \forall\ t \in [0,\tau] \tag{11.6}$$

with initial condition $B(0) = 1$. The value of this money market account must satisfy:

$$0 < B(\omega, t) < +\infty \qquad \text{a.e. } Q \qquad \forall\ t \in [0,\tau]$$

To guarantee that this condition is satisfied, we require:

$$\int_0^\tau |f(0,v)|\, dv < +\infty \quad \text{and} \quad \int_0^\tau \left(\int_0^t |\alpha(\omega, v, t)|\, dv\right) dt < +\infty \qquad \text{a.e. } Q$$

11.2.2. Bond price process. Here technical conditions are applied to parameters of the bond price process, thereby allowing the resulting bond price process to be well behaved.

CONDITION 3. *Regularity of the bond price process.* To ensure a well-behaved bond price process, the following regularity conditions are imposed:

$$\int_0^t \left(\int_v^t \sigma_i(\omega, v, y)\, dy\right)^2 dv < +\infty \qquad \text{a.e. } Q \qquad \forall\ t \in [0,\tau],\ i = 1, \ldots, n$$

$$\int_0^t \left(\int_T^t \sigma_i(\omega, v, y)\, dy\right)^2 dv < +\infty \qquad \text{a.e. } Q \qquad \forall\ T \in [0,\tau],\ t \in [0,T],\ i = 1, \ldots, n$$

and

$$t \to \int_t^T \left(\int_0^t \sigma_i(\omega, v, y)\, dz_i(v)\right)^2 dy \qquad \text{is continuous a.e. } Q \qquad \forall\ T \in [0,\tau],\ i = 1, \ldots, n$$

Given Conditions 2 and 3 and using the lemma and two corollaries below, we determine the bond price process.

LEMMA 0.1. This is a generalised form of the Fubini theorem for stochastic integrals. Given the following:

$$\begin{aligned} (\Omega, F, Q) \quad &- \quad \text{probability space,} \\ \{F_t\} \quad &- \quad \text{filtration generated by a Brownian motion } \{z(t) : t \in [0, \tau]\} \end{aligned}$$

let $\{\Phi(\omega, t, a) : (t, a) \in [0, \tau] \times [0, \tau]\}$ be a family of real random variables such that

(i) $((\omega, t), a) \in \{(\Omega \times [0, \tau]) \times [0, \tau]\} \to \Phi(\omega, t, a)$ is $L \times B[0, \tau]$ measurable[3],

(ii) $\displaystyle\int_0^t \Phi^2(\omega, s, a)\, ds < +\infty \quad \text{a.e.} \quad \forall\ t \in [0, \tau];$

(iii) $\displaystyle\int_0^t \left(\int_0^\tau \Phi(\omega, s, a)\, da \right)^2 ds < +\infty \quad \text{a.e.} \quad \forall\ t \in [0, \tau].$

If $t \to \int_0^\tau \Big(\int_0^t \Phi(\omega, s, a)\, dz(s) \Big)\, da$ is continuous a.e. then:

$$\int_0^t \left(\int_0^\tau \Phi(\omega, s, a)\, da \right) dz(s) = \int_0^\tau \left(\int_0^t \Phi(\omega, s, a)\, dz(s) \right) da \qquad \forall\ t \in [0, \tau]$$

PROOF. Let χ_A and χ_B be characteristic functions such that:

$$\chi_A(\omega) = \begin{cases} 1 & \text{for } \omega \in A \\ 0 & \text{for } \omega \in A^C \end{cases}$$

$$\text{and} \qquad \chi_B(\omega) = \begin{cases} 1 & \text{for } \omega \in B \\ 0 & \text{for } \omega \in B^C \end{cases}$$

where A is a set $\{t : t \in [s, \tau)\}$ and $B \in F_s$. Now we have:

$$\int_B \int_A \chi_A \chi_B\, dt\, dz = \int_B \lambda(A)\, \chi_B\, dz = \lambda(A)\, Q(B)$$

where λ is the Lebesque measure and Q the measure associated with filtration F. Also:

$$\int_A \int_B \chi_A \chi_B\, dz\, dt = \int_A \chi_A\, Q(B) dt = \lambda(A)\, Q(B)$$

[3] L is the smallest σ-field on $(\Omega \times [0, \tau])$ such that all left-continuous F_t-adapted processes $Y : (\omega, t) \in (\Omega \times [0, \tau]) \to Y(\omega, t) \in R^d$ are measurable.

Therefore:

$$\int_B \int_A \chi_A \chi_B \, dt \, dz = \int_A \int_B \chi_A \chi_B \, dz \, dt \tag{11.7}$$

Let $\mathcal{V}$ be a class of functions $h(t,\omega) : [0,\infty] \times \Omega \to R$ such that:

- $(t,\omega) \to h(t,\omega)$ is $\mathcal{B} \times F-$ measurable[4],
- for each $t \geq 0$ the function $\omega \to h(t,\omega)$ is F_t-measurable and
- $\mathbb{E}\left[\int_S^T h^2(t,\omega)\, dt\right] < \infty$

Now an elementary function $\psi \in \mathcal{V}$ may be defined as a sum of characteristic function as[5]:

$$\psi(t,\omega) = \sum_{j \geq 0} e_j(\omega) \cdot \chi_{[t_j, t_{j+1})}(t)$$

Since $\phi \in \mathcal{V}$, each function e_j must be F_{t_j}-measurable. Hence we may define:

$$\int_S^T \psi(t,\omega)\, dz_t(\omega) = \sum_{j \geq 0} e_j(\omega)\left[z_{t_{j+1}} - z_{t_j}\right](\omega)$$

for some $0 \leq S \leq T$. Therefore the Ito Integral of some function $h \in \mathcal{V}$ may be written as:

$$\int_S^T h(t,\omega)\, dz_t(\omega) = \lim_{n\to\infty} \int_S^T \psi_n(t,\omega)\, dz_t(\omega)$$

where the limit is taken in $L^2(P)$ and $\{\psi_n\}$ is a sequence of elementary functions such that:

$$\mathbb{E}\left[\int_S^T \big(h(t,\omega) - \psi_n(t,\omega)\big)^2 \, dt\right] \to 0 \quad \text{as} \quad n \to \infty$$

We have shown that the integral of any function $h \in \mathcal{V}$ may be written as the limit of the integral of a sequence of elementary functions. The elementary functions may be expressed as sums of characteristic functions. A similar results (used in the proof of the standard Fubini Theorem, e.g. [**16**]) exists for the purely deterministic case. Hence we conclude that since relationship (11.7) holds for characteristic functions, it holds for any function $\Phi(\omega, t, a)$.

For an alternative description of this proof see [**33**, Chapter 3, Problem 6.12].

□

Corollary 0.1.1. Assume Lemma 0.1 holds and define:

$$\Phi(\omega, s, a) = \begin{cases} 0 & \text{if } (s,a) \notin [0,t] \times [t,\tau] \\ \sigma(\omega, s, a) & \text{if } (s,a) \in [0,t] \times [t,\tau] \end{cases}$$

---

[4] As before $\mathcal{B}$ represents the Borel σ-algebra on $[0,\infty)$.

[5] The following has been adapted from [**43**].

Then

$$\int_0^y \left(\int_t^\tau \sigma(\omega, s, a)\, da \right) \, dz(s) = \int_t^\tau \left(\int_0^y \sigma(\omega, s, a)\, dz(s) \right) \, da \qquad \forall \;\; y \in [0, t]$$

PROOF.

$$\begin{aligned}
&\int_0^y \int_t^\tau \sigma(\omega, s, a)\, da\, dz(s) \\
&= \int_0^y \int_0^\tau \Phi(\omega, s, a)\, da\, dz(s) \\
&= \int_0^\tau \int_0^y \Phi(\omega, s, a)\, dz(s)\, da \qquad \text{by Lemma 0.1} \\
&= \int_t^\tau \int_0^y \sigma(\omega, s, a)\, dz(s)\, da
\end{aligned}$$

□

COROLLARY 0.1.2. Assume Lemma 0.1 holds and define:

$$\Phi(\omega, s, a) = \begin{cases} 0 & \text{if } (s, a) \notin [0, t] \times [0, t] \\ \sigma(\omega, s, a) 1_{s \le a} & \text{if } (s, a) \in [0, t] \times [0, t] \end{cases}$$

Then

$$\int_0^y \left(\int_s^t \sigma(\omega, s, a)\, da \right) \, dz(s) = \int_0^t \left(\int_0^{a \wedge y} \sigma(\omega, s, a)\, dz(s) \right) \, da \qquad \forall \;\; y \in [0, t]$$

PROOF.

$$\begin{aligned}
&\int_0^y \int_s^t \sigma(\omega, s, a)\, da\, dz(s) \\
&= \int_0^y \int_0^t \Phi(\omega, s, a)\, da\, dz(s) \qquad \text{since } \Phi = 0 \text{ for } a < s \\
&= \int_0^t \int_0^y \Phi(\omega, s, a)\, dz(s)\, da \qquad \text{by Lemma 0.1} \\
&= \int_0^t \int_0^{a \wedge y} \sigma(\omega, s, a)\, dz(s)\, da \qquad \text{since } \Phi = 0 \text{ for } s > a
\end{aligned}$$

□

Now consider the bond price (11.2):

$$\begin{aligned}
P(t, T) &= \exp\left(- \int_t^T f(t, y)\, dy \right) \\
\Rightarrow \ln P(t, T) &= - \int_t^T f(t, y)\, dy
\end{aligned}$$

Substituting (11.4) we have[6]:

$$\ln P(t,T) = -\int_t^T f(0,y)\,dy - \int_t^T\left(\int_0^t \alpha(v,y)\,dv\right)\,dy - \sum_{i=1}^n \int_t^T\left(\int_0^t \sigma_i(v,y)\,dz_i(v)\right)\,dy \tag{11.8}$$

Now, apply the standard Fubini theorem to the double integral on $\alpha(v,y)$ and Corollary 0.1.1 to the double integral on $\sigma_i(v,y)$ to get:

$$\begin{aligned}
\ln P(t,T) &= -\int_t^T f(0,y)\,dy - \int_0^t\left(\int_t^T \alpha(v,y)\,dy\right)\,dv \\
&\quad -\sum_{i=1}^n \int_0^t\left(\int_t^T \sigma_i(v,y)\,dy\right)\,dz_i(v) \\
&= -\int_0^T f(0,y)\,dy - \int_0^t\left(\int_v^T \alpha(v,y)\,dy\right)\,dv \\
&\quad -\sum_{i=1}^n \int_0^t\left(\int_v^T \sigma_i(v,y)\,dy\right)\,dz_i(v) \\
&\quad +\int_0^t f(0,y)\,dy + \int_0^t\left(\int_v^t \alpha(v,y)\,dy\right)\,dv \\
&\quad +\sum_{i=1}^n \int_0^t\left(\int_v^t \sigma_i(v,y)\,dy\right)\,dz_i(v)
\end{aligned} \tag{11.9}$$

Applying Corollary 0.1.2 to the last two terms of (11.9) gives:

$$\begin{aligned}
&\int_0^t\left(\int_v^t \alpha(v,y)\,dy\right)\,dv + \sum_{i=1}^n \int_0^t\left(\int_v^t \sigma_i(v,y)\,dy\right)\,dz_i(v) \\
&= \int_0^t\left(\int_0^{y\wedge t} \alpha(v,y)\,dv\right)\,dy + \sum_{i=1}^n \int_0^t\left(\int_0^{y\wedge t} \sigma_i(v,y)\,dz_i(v)\right)\,dy
\end{aligned}$$

and from (11.2) we know:

$$-\int_0^T f(0,y)\,dy = \ln P(0,T)$$

[6]Here to improve readability, we suppress the notational dependence on ω.

Hence (11.9) becomes:

$$\begin{aligned}
\ln P(t,T) &= \ln P(0,T) - \int_0^t \left(\int_v^T \alpha(v,y)\,dy \right) dv \\
&\quad - \sum_{i=1}^n \int_0^t \left(\int_v^T \sigma_i(v,y)\,dy \right) dz_i(v) \\
&\quad + \int_0^t f(0,y)\,dy + \int_0^t \left(\int_0^{y\wedge t} \alpha(v,y)\,dv \right) dy \\
&\quad + \sum_{i=1}^n \int_0^t \left(\int_0^{y\wedge t} \sigma_i(v,y)\,dz_i(v) \right) dy
\end{aligned}$$

However, from (11.5) we have[7]:

$$\begin{aligned}
\int_0^t r(v)\,dv &= \int_0^t f(0,y)\,dy + \int_0^t \left(\int_0^{y\wedge t} \alpha(v,y)\,dv \right) dy \\
&\quad + \sum_{i=1}^n \int_0^t \left(\int_0^{y\wedge t} \sigma_i(v,y)\,dz_i(v) \right) dy
\end{aligned}$$

and so the dynamics of the bond price process are:

$$\begin{aligned}
\ln P(t,T) &= \ln P(0,T) + \int_0^t r(y)\,dy - \int_0^t \left(\int_v^T \alpha(v,y)\,dy \right) dv \\
&\quad - \sum_{i=1}^n \int_0^t \left(\int_v^T \sigma_i(v,y)\,dy \right) dz_i(v)
\end{aligned} \tag{11.10}$$

Let:

$$a_i(\omega,t,T) = -\int_t^T \sigma_i(\omega,t,y)\,dy \qquad \text{for } i = 1,\dots,n \tag{11.11a}$$

$$b(\omega,t,T) = -\int_t^T \alpha(\omega,t,y)\,dy + \frac{1}{2}\sum_{i=1}^n a_i^2(\omega,t,T) \tag{11.11b}$$

[7] Directly integrating (11.5) yields:

$$\int_0^t r(y)\,dy = \int_0^t f(0,y)\,dy + \int_0^t \left(\int_0^t \alpha(v,y)\,dv \right) dy + \sum_{i=1}^n \int_0^t \left(\int_0^t \sigma_i(v,y)\,dz_i(v) \right) dy$$

However, by definition of $\alpha(v,y)$ and $\sigma_i(v,y)$ as drift and volatility parameters of the forward rate process, we require $v \le y$ for all $v, y \in [0,\tau]$ and hence the upper limit on the inner integrals must become $\min(y,t) \equiv y \wedge t$.

and (11.10) becomes:

$$\begin{aligned}(11.12)\quad \ln P(t,T) &= \ln P(0,T) + \int_0^t \big(r(v) + b(v,T)\big)\,dv \\ &\quad - \frac{1}{2}\sum_{i=1}^n \int_0^t a_i^2(v,T)\,dv + \sum_{i=1}^n \int_0^t a_i(v,T)\,dz_i(v) \ \text{ a.e. } Q\end{aligned}$$

which may be expressed in differential form as:

$$d\ln P(t,T) = \left(r(t) + b(t,T) - \frac{1}{2}\sum_{i=1}^n a_i^2(t,T)\right)dt + \sum_{i=1}^n a_i(t,T)\,dz_i(t)$$

Now applying Ito's Lemma, we find the differential equation satisfied by the bond price $P(t,T)$, to be:

$$dP(t,T) = \frac{\partial P}{\partial(\ln P)}\,d(\ln P) + \frac{1}{2}\frac{\partial^2 P}{\partial(\ln P)^2}\,d(\ln P)\,d(\ln P) + \frac{\partial P}{\partial t}\,dt$$

where

$$\frac{\partial P}{\partial(\ln P)} = P = \frac{\partial^2 P}{\partial(\ln P)^2}$$

and so:

$$\begin{aligned} dP(t,T) &= P(t,T)\left(r(t) + b(t,T) - \frac{1}{2}\sum_{i=1}^n a_i^2(t,T)\right)dt \\ &\quad + P(t,T)\sum_{i=1}^n a_i(t,T)\,dz_i(t) + \frac{1}{2}P(t,T)\sum_{i=1}^n a_i^2(t,T)\,dt \\ (11.13)\qquad &= P(t,T)\left(r(t) + b(t,T)\right)\,dt + P(t,T)\sum_{i=1}^n a_i(t,T)\,dz_i(t) \qquad \text{a.e. } Q\end{aligned}$$

Since in (11.13) both the drift term $r(t) + b(\omega,t,T)$ and the volatility coefficients
$a_i(\omega,t,T),\ i = 1,\ldots,n$, may depend on the history of the Brownian motions, the bond price process is non-Markovian.

11.2.3. Relative bond price process. Let $Z(t,T) = \frac{P(t,T)}{B(t)};\ T \in [0,\tau],\ t \in [0,T]$ be the time t relative price of a T-maturity bond. Here, the bond price is expressed in terms of the money market account, so its drift with respect to the short-term interest rate is removed. Make use of Ito's Lemma to determine the dynamics of the relative bond price as[8]:

[8]Dynamics of the money market account are easily found from (11.6) as:

$$\begin{aligned}
dZ(t,T) &= \frac{\partial Z(t,T)}{\partial B(t)}\,dB(t) + \frac{\partial Z(t,T)}{\partial P(t,T)}\,dP(t,T) \\
&\quad + \frac{1}{2}\frac{\partial^2 Z(t,T)}{\partial P(t,T)^2}\,dP(t,T)\,dP(t,T) \\
&= -\frac{P(t,T)}{B(t)^2}\,dB(t) + \frac{dP(t,T)}{B(t)} + 0 \\
&= -\frac{P(t,T)B(t)r(t)}{B(t)^2}\,dt + \big(r(t)+b(t,T)\big)\frac{P(t,T)}{B(t)}\,dt \\
&\quad + \sum_{i=1}^{n} a_i(t,T)\frac{P(t,T)}{B(t)}\,dz_i(t) \\
&= Z(t,T)\,b(t,T)\,dt + Z(t,T)\sum_{i=1}^{n} a_i(t,T)\,dz_i(t) \qquad (11.14)
\end{aligned}$$

Also:

$$\begin{aligned}
d\ln Z(t,T) &= \frac{\partial \ln Z(t,T)}{\partial Z(t,T)}\,dZ(t,T) + \frac{1}{2}\frac{\partial^2 \ln Z(t,T)}{\partial Z(t,T)^2}\,dZ(t,T)\,dZ(t,T) \\
&= \frac{1}{Z(t,T)}\,dZ(t,T) - \frac{1}{2}\frac{1}{Z^2(t,T)}Z^2(t,T)\sum_{i=1}^{n} a_i^2(t,T)\,dt \\
&= \left(b(t,T) - \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T)\right)dt + \sum_{i=1}^{n} a_i(t,T)\,dz_i(t) \qquad (11.15)
\end{aligned}$$

Hence the integral form of the relative bond price process is:

$$\begin{aligned}
\ln Z(t,T) &= \ln Z(0,T) + \int_0^t b(v,T)\,dv - \frac{1}{2}\sum_{i=1}^{n}\int_0^t a_i^2(v,T)\,dv \\
&\quad + \sum_{i=1}^{n}\int_0^t a_i(v,T)\,dz_i(v) \qquad \text{a.e. } Q \qquad (11.16)
\end{aligned}$$

Again, the relative bond price is non-Markovian since the drift and volatility coefficients may depend on the history of the Brownian motions through the cumulative forward rate drift and volatility terms $b(\omega,\cdot,T)$ and $a_i(\omega,\cdot,T),\ i = 1,\ldots,n$.

$$dB(t) = \exp\Big(\int_0^t r(y)\,dy\Big) r(t)dt = B(t)r(t)dt$$

11.3. Arbitrage-free framework

In (11.13) we found the bond price process, under the probability measure Q, to be:

$$dP(t,T) = P(t,T)\,(r(t) + b(t,T))\;dt + P(t,T)\sum_{i=1}^{n} a_i(t,T)\,dz_i(t)$$

Absence of arbitrage may be equated to the existence of a probability measure $\widetilde{Q}$ such that all discounted security prices are martingales [**22**]. We wish to find this martingale probability measure such that it is equivalent[9] to the market measure Q and the drift in (11.13) becomes $r(t)$. Hence we wish to find a vector of Brownian motions $\{\tilde{z}_1(t),\ldots,\tilde{z}_n(t)\,:\,t\in[0,\tau]\}$ with $n\geq 1$ (and associated adapted processes $\{\gamma_i(t);t\in[0,\tau]\}$ which define $\tilde{z}_i(t) = -\int_0^t \gamma_i(u)\,du + z_i(t),\ i = 1,\ldots,n$) such that[10]:

$$\begin{aligned} dP(t,T) &= P(t,T)\Big(r(t) + b(t,T) + \sum_{i=1}^{n} a_i(t,T)\gamma_i(t)\Big)\,dt \\ &\quad + P(t,T)\sum_{i=1}^{n} a_i(t,T)\ (-\gamma_i(t)dt + dz_i(t)) \end{aligned}$$

with

$$b(t,T) + \sum_{i=1}^{n} a_i(t,T)\gamma_i(t) = 0 \tag{11.17}$$

Since, by definition, $b(t,T)$ represents the cumulative drift of forward rates over $[t,T]$, (11.17) implies a restriction is required on the forward rate process to ensure arbitrage-free pricing.

11.3.1. Specification of the martingale measure. Let us examine necessary and sufficient conditions on the forward rate process ensuring the existence of a unique, equivalent martingale measure and hence arbitrage-free pricing.

---

[9] Equivalence of two probability measures, Q and $\widetilde{Q}$ implies $Q(A) = 0$ if and only if $\widetilde{Q}(A) = 0,\ A\in F$, that is, the probability measures have the same null sets [**24**].

[10] This is equivalent to finding an equivalent probability measure such that the relative bond price process (11.14) has no drift term, i.e. may be expressed as:

$$dZ(t,T) = Z(t,T)\sum_{i=1}^{n} a_i(t,T)\,d\tilde{z}_i(t)$$

Condition 4. *Existence of the market prices of risk.* Set $S_1, \dots, S_n \in [0, \tau]$ such that $0 < S_1 < S_2 < \dots < S_n \le \tau$ and let λ be a Lebesque measure. Assume there exist solutions

$$\gamma_i(\cdot, \cdot;\ S_1, \dots, S_n) : \Omega \times [0, S_1] \to R \quad \text{for } i = 1, \dots, n \quad \text{a.e. } Q \times \lambda$$

to a system of equations specified as:

$$\begin{bmatrix} b(t, S_1) \\ \vdots \\ b(t, S_n) \end{bmatrix} + \begin{bmatrix} a_1(t, S_1) & \dots & a_n(t, S_1) \\ & \vdots & \\ a_1(t, S_n) & \dots & a_n(t, S_n) \end{bmatrix} \begin{bmatrix} \gamma_1(t; S_1, \dots, S_n) \\ \vdots \\ \gamma_n(t; S_1, \dots, S_n) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{11.18}$$

Also assume that these solutions satisfy the following three conditions[11]:

$$\int_0^{S_1} \gamma_i^2(v; S_1, \dots, S_n)\, dv < +\infty \qquad \text{a.e. } Q \ \text{ for } i = 1, \dots, n \tag{11.19a}$$

$$\mathbb{E}^Q \left[\exp\left(\sum_{i=1}^n \int_0^{S_1} \gamma_i(v; S_1, \dots, S_n) dz_i(v) - \frac{1}{2} \sum_{i=1}^n \int_0^{S_1} \gamma_i^2(v; S_1, \dots, S_n) dv \right) \right] = 1 \tag{11.19b}$$

$$\mathbb{E}^Q \left[\exp\left(\sum_{i=1}^n \int_0^{S_1} \big(a_i(v, y) + \gamma_i(v; S_1, \dots, S_n)\big)\, dz_i(v) - \frac{1}{2} \sum_{i=1}^n \int_0^{S_1} \big(a_i(v, y) + \gamma_i(v; S_1, \dots, S_n)\big)^2 dv \right) \right] = 1 \tag{11.19c}$$

for $y \in \{S_1, \dots, S_n\}$

From the system of equations (11.18), we may view $\gamma_i(t; S_1, \dots, S_n),\ i = 1, \dots, n$, as the market prices of risk associated with each source of uncertainty, that is with each Brownian motion $z_i(t),\ i = 1, \dots, n$.

The market price of risk is defined as the mean rate of return in excess of the risk-free rate of interest, normalised by the volatility of that return. From (11.13) the instantaneous expected return on a T-maturity bond is $r(t) + b(t, T)$ with $a_i(t, T),\ i = 1, \dots, n$, the corresponding instantaneous standard deviations of bond return due to the i^{th} random factor. Therefore, interpreting $\gamma_i(t; S_1, \dots, S_n),\ i = 1, \dots, n$, as the market prices of risk associated with the sources of uncertainty, this market price of risk relationship may be expressed as:

[11] $\mathbb{E}^P[\cdot]$ is the expectation taken with respect to some probability measure P.

$$b(t,T) = \sum_{i=1}^{n} a_i(t,T)\,(-\gamma_i(t;S_1,\ldots,S_n)) \tag{11.20}$$

where the market prices of risk $\gamma_i(t;S_1,\ldots,S_n)$, $i=1,\ldots,n$, depend on the specific choice of bond maturities $\{S_1,\ldots,S_n\}$. We see that (11.20) is equivalent to (11.17) which was derived using a heuristic argument[12].
The above Condition 4 guarantees the existence of an equivalent martingale measure. This is shown by means of the following proposition.

PROPOSITION 1. *Existence of an equivalent martingale probability measure. Set* $S_1,\ldots,S_n \in [0,\tau]$ *such that* $0 < S_1 < S_2 < \ldots < S_n \le \tau$. *Given a vector of forward rate drifts* $\{\alpha(\cdot,S_1),\ldots,\alpha(\cdot,S_n)\}$ *and volatilities* $\{\sigma_i(\cdot,S_1),\ldots,\sigma_i(\cdot,S_n)\}$, $i=1,\ldots,n$, *which satisfy Conditions 1 - 3, then Condition 4 holds if and only if there exists an equivalent probability measure* $\tilde{Q}_{S_1,\ldots,S_n}$ *such that* $\{Z(t,S_1),\ldots,Z(t,S_n)\}$ *are martingales with respect to* $\{F_t : t\in[0,S_1]\}$.

PROOF. The proof of this proposition is by the following two lemmas.

LEMMA 1.1. Assume Conditions 1–3 hold for some fixed $\{S_1,\ldots,S_n\}\in[0,\tau]$; $0<S_1<S_2<\ldots<S_n\le\tau$.
Define

$$X(t,y) = \int_0^t b(v,y)\,dv + \sum_{i=1}^{n}\int_0^t a_i(v,y)\,dz_i(v) \quad \forall\, t\in[0,y],\ y\in\{S_1,\ldots,S_n\}$$

Then $\gamma_i : \Omega\times[0,\tau]\to R$ for $i=1,\ldots,n$, satisfies the following four conditions:

$$\text{(i)}\quad \begin{bmatrix} b(t,S_1)\\ \vdots \\ b(t,S_n)\end{bmatrix} + \begin{bmatrix} a_1(t,S_1) & \ldots & a_n(t,S_1)\\ & \vdots & \\ a_1(t,S_n) & \ldots & a_n(t,S_n)\end{bmatrix}\begin{bmatrix}\gamma_1(t)\\ \vdots\\ \gamma_n(t)\end{bmatrix} = \begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix} \quad \text{a.e. } Q\times\lambda$$

$$\text{(ii)}\quad \int_0^{S_1}\gamma_i^2(v)\,dv < +\infty \qquad \text{a.e. } Q \text{ for } i=1,\ldots,n$$

$$\text{(iii)}\quad \mathbb{E}^Q\left[\exp\left(\sum_{i=1}^{n}\int_0^{S_1}\gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1}\gamma_i^2(v)\,dv\right)\right] = 1$$

[12]The two formulae are equivalent up to the former's dependence on a specific choice of bond maturities. However, later in the analysis this dependence is eliminated.

(iv) $$\mathbb{E}^Q\left[\exp\left(\sum_{i=1}^{n}\int_0^{S_1}(a_i(v,y)+\gamma_i(v))\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1}(a_i(v,y)+\gamma_i(v))^2\,dv\right)\right] = 1 \quad \text{for } y\in\{S_1,\dots,S_n\}$$

if and only if there exists a probability measure $\widetilde{Q}_{S_1,\dots,S_n}$ such that the following are true:

(a) $$\frac{d\widetilde{Q}_{S_1,\dots,S_n}}{dQ} = \exp\left(\sum_{i=1}^{n}\int_0^{S_1}\gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1}\gamma_i^2(v)\,dv\right)$$

(b) $\tilde{z}_i^{S_1,\dots,S_n}(t) = z_i(t) - \int_0^t \gamma_i(v)\,dv$ are Brownian motions on $\left\{\left(\Omega,F,\widetilde{Q}_{S_1,\dots,S_n}\right),\left(F_t; t\in[0,S_1]\right)\right\}$ for $i=1,\dots,n$

(c) $$\begin{bmatrix} dX(t,S_1)\\ \vdots \\ dX(t,S_n)\end{bmatrix} = \begin{bmatrix} a_1(t,S_1) & \dots & a_n(t,S_1)\\ & \vdots & \\ a_1(t,S_n) & \dots & a_n(t,S_n)\end{bmatrix}\begin{bmatrix} d\tilde{z}_1^{S_1,\dots,S_n}(t)\\ \vdots \\ d\tilde{z}_n^{S_1,\dots,S_n}(t)\end{bmatrix}$$
for $t\in[0,S_1]$

(d) $Z(t,S_i)$ are martingales on $\left\{\left(\Omega,F,\widetilde{Q}_{S_1,\dots,S_n}\right),\left(F_t; t\in[0,S_1]\right)\right\}$ for $i=1,\dots,n$

PROOF. In the first part of this proof assume (i)–(iv) hold and show that (a)–(d) hold. Make use of the Girsanov Theorem; see [**21**, Corollary 13.25]. Properties (ii) and (iii) satisfy the conditions for the Girsanov Theorem, so there exists a probability measure $\widetilde{Q}_{S_1,\dots,S_n}$ on (Ω,F) defined by (a) and a corresponding n-dimensional Brownian motion on $(\Omega,F,\widetilde{Q}_{S_1,\dots,S_n})$ defined by (b).

Consider the stochastic differential equation (11.14) describing the evolution of the relative bond price $Z(t,y)$:

$$dZ(t,y) = Z(t,y)\,b(t,y)\,dt + Z(t,y)\sum_{i=1}^{n}a_i(t,y)\,dz_i(t)$$
$$\frac{dZ(t,y)}{Z(t,y)} = b(t,y)\,dt + \sum_{i=1}^{n}a_i(t,y)\,dz_i(t)$$

hence

$$\int_0^t\frac{dZ(v,y)}{Z(v,y)} = \int_0^t b(v,y)\,dv + \sum_{i=1}^{n}\int_0^t a_i(v,y)\,dz_i(v) = X(t,y) \tag{11.21}$$

by the definition of $X(t,y)$ and for all $t\in[0,y]$, $y\in\{S_1,\dots,S_n\}$. Therefore:

$$dX(t,y) = b(t,y)\,dt + \sum_{i=1}^{n} a_i(t,y)\,dz_i(t)$$

and from (i) we have:

$$b(t,y) = \sum_{i=1}^{n} a_i(t,y)\big(-\gamma_i(t)\big)$$

so:

$$\begin{aligned} dX(t,y) &= \sum_{i=1}^{n} a_i(t,y)\big(-\gamma_i(t)\big)\,dt + \sum_{i=1}^{n} a_i(t,y)\,dz_i(t) \\ &= \sum_{i=1}^{n} a_i(t,y)\big(dz_i(t) - \gamma_i(t)\,dt\big) \\ &= \sum_{i=1}^{n} a_i(t,y)\,d\tilde{z}_i^{S_1,\dots,S_n}(t) \end{aligned}$$

by definition of $\tilde{z}_i^{S_1,\dots,S_n}$ in (b) and for all $t \in [0,S_1]$, $y \in \{S_1,\dots,S_n\}$. Hence (c) follows.

From (11.21) and (c) we have:

$$\begin{aligned} \frac{dZ(t,y)}{Z(t,y)} &= dX(t,y) = \sum_{i=1}^{n} a_i(t,y)\,d\tilde{z}_i^{S_1,\dots,S_n}(t) \\ \Rightarrow Z(t,y) &= Z(0,y) + \sum_{i=1}^{n} \int_0^t Z(v,y)\,a_i(v,y)\,d\tilde{z}_i^{S_1,\dots,S_n}(v) \end{aligned}$$

Now by [**43**, Corollary 3.2.6] $Z(t,y)$ is a martingale and hence (d) is satisfied.

Conversely, assume (a)–(d) hold and show that (i)–(iv) are true.

Again by (11.21) and (c) we have:

$$\frac{dZ(t,y)}{Z(t,y)} = dX(t,y) = \sum_{i=1}^{n} a_i(t,y)\,d\tilde{z}_i^{S_1,\dots,S_n}(t)$$

where $d\tilde{z}_i^{S_1,\dots,S_n}(t)$, $i = 1,\dots,n$, are defined in (b) as:

$$d\tilde{z}_i^{S_1,\dots,S_n}(t) = dz_i(t) - \gamma_i(t)\,dt$$

Hence:

$$\begin{aligned}\frac{dZ(t,y)}{Z(t,y)} &= \sum_{i=1}^{n} a_i(t,y)\left(dz_i(t) - \gamma_i(t)\,dt\right)\\ &= \sum_{i=1}^{n} a_i(t,y)\left(-\gamma_i(t)\right)dt + \sum_{i=1}^{n} a_i(t,y)\,dz_i(t)\end{aligned}$$

But in (11.14) the process for $Z(t,y)$ is defined as:

$$dZ(t,T) = Z(t,T)\,b(t,T)\,dt + Z(t,T)\sum_{i=1}^{n} a_i(t,T)\,dz_i(t)$$

Hence:

$$b(t,y) = \sum_{i=1}^{n} a_i(t,y)\left(-\gamma_i(t)\right)$$

and (i) holds. The stochastic integral of any variable $\gamma_i(t)$ with respect to a Brownian motion $z_i(t)$ is defined only if $\gamma_i(t)$ is square integrable (see [**30**, Definition 1.5]). Since this integral is defined in (a), $\gamma_i(t)$ is square integrable, and hence (ii) holds. The Radon–Nikodym derivative (a) defines the relationship between measure Q and $\widetilde{Q}_{S_1,\ldots,S_n}$; hence:

$$\begin{aligned}1 &= \mathbb{E}^{\widetilde{Q}_{S_1,\ldots,S_n}}\left[1\right]\\ &= \mathbb{E}^{Q}\left[\exp\left(\sum_{i=1}^{n}\int_0^{S_1}\gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1}\gamma_i^2(v)\,dv\right)\right]\end{aligned}$$

which proves (iii). We have

$$Z(t,y) = Z(0,y)\exp\left(\sum_{i=1}^{n}\int_0^{t} a_i(v,y)\,d\tilde{z}^{S_1,\,\ldots\,,\,S_n}(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{t} a_i^2(v,y)\,dv\right)$$

and by (d) $Z(t,y)$, $t \in [0,S_1]$, $y \in \{S_1,\ldots S_n\}$ is a martingale, so:

$$\begin{aligned}Z(0,y) &= \mathbb{E}^{\widetilde{Q}_{S_1,\ldots,S_n}}\left[Z(S_1,y)\right]\\ &= \mathbb{E}^{Q}\Bigg[Z(0,y)\exp\left(\sum_{i=1}^{n}\int_0^{S_1} a_i(v,y)\,d\tilde{z}^{S_1,\,\ldots\,,\,S_n}(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1} a_i^2(v,y)\,dv\right)\\ &\qquad \times \exp\left(\sum_{i=1}^{n}\int_0^{S_1}\gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1}\gamma_i^2(v)\,dv\right)\Bigg]\end{aligned}$$

$$= \mathbb{E}^Q \Bigg[Z(0,y) \exp\left(\sum_{i=1}^n \int_0^{S_1} a_i(v,y)\, dz_i(v) - \sum_{i=1}^n \int_0^{S_1} a_i(v,y)\gamma_i(v)\, dv \right)$$
$$\times \exp\Bigg(-\frac{1}{2} \sum_{i=1}^n \int_0^{S_1} a_i^2(v,y)\, dv + \sum_{i=1}^n \int_0^{S_1} \gamma_i(v)\, dz_i(v)$$
$$-\frac{1}{2} \sum_{i=1}^n \int_0^{S_1} \gamma_i^2(v)\, dv \Bigg) \Bigg]$$
$$= Z(0,y) \mathbb{E}^Q \Bigg[\exp\Bigg(\sum_{i=1}^n \int_0^{S_1} \big(a_i(v,y) + \gamma_i(v)\big)\, dz_i(v)$$
$$-\frac{1}{2} \sum_{i=1}^n \int_0^{S_1} \big(a_i(v,y) + \gamma_i(v)\big)^2\, dv \Bigg) \Bigg]$$

Hence (iv) holds. □

Lemma 1.2. Assume Conditions 1–3 hold for some fixed $\{S_1, \ldots, S_n\} \in [0, \tau]$; $0 < S_1 < S_2 < \ldots < S_n \leq \tau$.
Define

$$X(t,y) = \int_0^t b(v,y)\, dv + \sum_{i=1}^n \int_0^t a_i(v,y)\, dz_i(v) \quad \forall\, t \in [0,y],\ y \in \{S_1, \ldots, S_n\}$$

Then there exists a probability measure $\overline{Q}$ equivalent to Q such that $Z(t, S_i)$ are martingales on $\left\{ \left(\Omega, F, \overline{Q}\right), \left(F_t; t \in [0, S_1]\right) \right\}$ for all $i = 1, \ldots, n$, if and only if there exists $\gamma_i : \Omega \times [0, \tau] \to R$ for $i = 1, \ldots, n$, and a probability measure $\widetilde{Q}_{S_1,\ldots,S_n}$ such that (a), (b), (c) and (d) of Lemma 1.1 hold.

Proof. First, suppose:

- there exists a probability measure $\overline{Q}$ equivalent to Q,
- $Z(t,y),\ y \in \{S_1, \ldots, S_n\}$ are martingales under $\overline{Q}$,

and we need to prove there exists:

- $\gamma_i,\ i = 1, \ldots, n$,
- probability measure $\widetilde{Q}_{S_1,\ldots,S_n}$,

such that (a)–(d) hold.

Since probability measure $\overline{Q}$ is equivalent to Q we know, by the Radon–Nikodym Theorem (see [**49**]), that there exists a non-negative random process $N(t),\ t \in [0, S_1]$ such that:

$$\frac{d\overline{Q}}{dQ} = N(S_1)$$

For any generic process to be a Radon–Nikodym derivative it must satisfy the following three conditions [**45**]:

(1) $N(0) = 1$,
(2) $N(t) > 0$ for all $t \in [0, S_1]$,
(3) $\mathbb{E}^Q\left[N(S_1)|F_t\right] = N(t)$.

Defining $N(t)$ such that:

$$N(t) = \exp\left(\sum_{i=1}^{n}\int_0^t \gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \gamma_i^2(v)\,dv\right)$$

then (1) and (2) are met. Taking the natural logarithm of $N(t)$ we have:

$$\begin{aligned}\ln N(t) &= \sum_{i=1}^{n}\int_0^t \gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \gamma_i^2(v)\,dv \\ \Rightarrow d\ln N(t) &= \sum_{i=1}^{n}\gamma_i(t)\,dz_i(t) - \frac{1}{2}\sum_{i=1}^{n}\gamma_i^2(t)\,dt\end{aligned}$$

By Ito's Lemma we have:

$$\begin{aligned}dN(t) &= \frac{\partial N(t)}{\partial \ln N(t)}\,d\ln N(t) + \frac{1}{2}\frac{\partial^2 N(t)}{\partial \ln N(t)^2}\,d\ln N(t)\,d\ln N(t) \\ &= N(t)\sum_{i=1}^{n}\gamma_i(t)\,dz_i(t) - \frac{1}{2}N(t)\sum_{i=1}^{n}\gamma_i^2(t)\,dt + \frac{1}{2}N(t)\sum_{i=1}^{n}\gamma_i^2(t)\,dt \\ &= N(t)\sum_{i=1}^{n}\gamma_i(t)\,dz_i(t)\end{aligned}$$

which, by the converse of the martingale representation theorem (see [**43**]), implies that $N(t)$ is a martingale and (3) holds. All the conditions for $N(t)$ to be a valid Radon–Nikodym derivative are met. Hence let $\widetilde{Q}_{S_1,\ldots,S_n}$ be the equivalent probability measure defined by this specific formulation of $N(t)$. Then (a) holds, and by the Girsanov Theorem (see [**43**, Theorem 8.6.4]) the new Brownian motion, $\tilde{z}_i^{S_1,\ldots,S_n}(t)$, $i = 1, \ldots, n$, corresponding to the measure $\widetilde{Q}_{S_1,\ldots,S_n}$ is defined by:

$$\tilde{z}_i^{S_1,\ldots,S_n}(t) = z_i(t) - \int_0^t \gamma_i(v)\,dv$$

and (b) holds.

$Z(t, y)$, $y \in \{S_1, \ldots, S_n\}$ are martingales under probability measure $\overline{Q}$, more specifically they are martingales under the specific probability measure $\widetilde{Q}_{S_1,\ldots,S_n}$, and (d) holds. By the martingale representation theorem we may write:

$$dZ(t, y) = \sum_{i=1}^{n}\varphi_i(t, y)\,d\tilde{z}_i^{S_1,\ldots,S_n}(t)$$

Now we may define $a_i(t,y) = \frac{\varphi_i(t,y)}{Z(t,y)}$, $i = 1, \dots, n$, and so:

$$dZ(t,y) = Z(t,y) \sum_{i=1}^{n} a_i(t,y)\, d\tilde{z}_i^{S_1,\,\dots\,,\,S_n}(t)$$

$$\Rightarrow \frac{dZ(t,y)}{Z(t,y)} = \sum_{i=1}^{n} a_i(t,y)\, d\tilde{z}_i^{S_1,\,\dots\,,\,S_n}(t)$$

However, by the definition of $X(t,y)$ we showed above in (11.21) that:

$$dX(t,y) = \frac{dZ(t,y)}{Z(t,y)} \qquad \text{for } y \in \{S_1, \dots, S_n\}$$

and hence (c) holds.

Conversely, suppose there exist

- γ_i, $i = 1, \dots, n$,
- probability measure $\widetilde{Q}_{S_1,\dots,S_n}$,

such that (a)–(d) are satisfied; we need to prove:

- there exists some probability measure $\overline{Q}$ equivalent to Q,
- $Z(t, S_i)$, $i = 1, \dots, n$, are martingales under $\overline{Q}$.

Since $\widetilde{Q}_{S_1,\dots,S_n}$ and Q are related by the Radon–Nikodym derivative specified in (a) they are equivalent, and we may set $\overline{Q} = \widetilde{Q}_{S_1,\dots,S_n}$. By (d) $Z(t, S_i)$, $i = 1, \dots, n$, are martingales under $\widetilde{Q}_{S_1,\dots,S_n}$ and since $\overline{Q} = \widetilde{Q}_{S_1,\dots,S_n}$ they are martingales under $\overline{Q}$.

□

Hence, by the proofs of Lemma 1.1 and Lemma 1.2 we have proved Proposition 1. □

This proposition states that, if Conditions 1–3 are satisfied, then Condition 4 is necessary and sufficient to allow an equivalent martingale measure $\widetilde{Q}_{S_1,\dots,S_n}$ to exist. An important feature of the proof of this proposition is Girsanov's Theorem, which defines the equivalent martingale measure as:

$$\frac{d\tilde{Q}_{S_1,\dots,S_n}}{dQ} = \exp\left(\sum_{i=1}^{n} \int_0^{S_1} \gamma_i(v; S_1, \dots, S_n) dz_i(v) - \frac{1}{2} \sum_{i=1}^{n} \int_0^{S_1} \gamma_i^2(v; S_1, \dots, S_n) dv \right) \tag{11.22}$$

and a new set of independent Brownian motions

$$\tilde{z}_i^{S_1,\,\dots\,,\,S_n}(t) = z_i(t) - \int_0^t \gamma_i(v; S_1, \dots, S_n)\, dv \quad \text{for } i = 1, \dots, n \tag{11.23}$$

on $\{(\Omega, F, \widetilde{Q}_{S_1,\dots,S_n}), \{F_t : t \in [0, S_1]\}\}$.

An additional constraint needs to be imposed to guarantee the uniqueness of the equivalent martingale measure.

CONDITION 5. *Uniqueness of the equivalent martingale probability measure.* Set $S_1, \ldots, S_n \in [0, \tau]$ such that $0 < S_1 < S_2 < \ldots < S_n \leq \tau$ and assume

$$\begin{bmatrix} a_1(t, S_1) & \ldots & a_n(t, S_1) \\ \vdots & & \vdots \\ a_1(t, S_n) & \ldots & a_n(t, S_n) \end{bmatrix}$$

is non-singular a.e. $Q \times \lambda$.

This Condition 5 is both necessary and sufficient to ensure that the equivalent martingale measure is unique. This is proved in the following proposition.

PROPOSITION 2. *Uniqueness of the equivalent martingale probability measure. Set $S_1, \ldots, S_n \in [0, \tau]$ such that $0 < S_1 < S_2 < \ldots < S_n \leq \tau$. Given a vector of forward rate drifts $\{\alpha(\cdot, S_1), \ldots, \alpha(\cdot, S_n)\}$ and volatilities $\{\sigma_i(\cdot, S_1), \ldots, \sigma_i(\cdot, S_n)\}$, $i = 1, \ldots, n$, which satisfy Conditions 1–4 then Condition 5 holds if and only if the martingale measure is unique.*

PROOF. The proof of this proposition is by the following two lemmas.

LEMMA 2.1. Set $S < \tau$ and define $\beta_i : \Omega \times [0, \tau] \to R$ for $i = 1, \ldots, n$, to be such that $\int_0^S \beta_i^2(v)\, dv < +\infty$ a.e. Q. Also define:

$$\begin{aligned} T_m &\equiv \inf\left\{ t \in [0, S];\ \mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^{n}\int_0^t \beta_i^2(v)\, dv\right)\right] \geq m \right\} \\ M^m(t) &\equiv \exp\left(\sum_{i=1}^{n}\int_0^{T_m \wedge t} \beta_i(v)\, dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m \wedge t} \beta_i^2(v)\, dv\right) \end{aligned}$$

Then

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\int_0^{S} \beta_i(v)\, dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S} \beta_i^2(v)\, dv\right)\right] = 1$$

if and only if $\{M^m(S)\}_{m=1}^{\infty}$ are uniformly integrable.

PROOF. Since $\beta_i(t)$, $i = 1, \ldots, n$, is by definition, a square integrable variable on (Ω, F, Q), then the Ito integral, given by $X^m(t) = \sum_{i=1}^{n}\int_0^{t \wedge T_m} \beta_i(v)\, dz_i(v)$ follows a continuous martingale on (Ω, F, Q) [**41**, Proposition B.1.2.]. Every martingale is a local martingale[13], so by [**21**, Lemma 13.17]

$$\begin{aligned} M^m(t) &= \exp\left(X^m(t) - \tfrac{1}{2}\langle X^m, X^m\rangle(t)\right) \\ &= \exp\left(\sum_{i=1}^n \int_0^{t\wedge T_m} \beta_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^n \int_0^{t\wedge T_m} \beta_i^2(v)\,dv\right) \end{aligned}$$

is a supermartingale.

By definition, T_m is the lowest $t \in [0,S]$ such that

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^n \int_0^t \beta_i^2(v)\,dv\right)\right] \geq m$$

Hence for all $t \leq T_m$ we have $\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^n \int_0^t \beta_i^2(v)\,dv\right)\right] \leq m$, and in particular $\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^n \int_0^{T_m} \beta_i^2(v)\,dv\right)\right] \leq m$. Hence

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^n \int_0^{S\wedge T_m} \beta_i^2(v)\,dv\right)\right] = \mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^n \int_0^{T_m} \beta_i^2(v)\,dv\right)\right] \leq m$$

and so by [**21**, Theorem 13.27] the supermartingale $M^m(t)$, $t \in [0,S]$ is a martingale and

$$\mathbb{E}\left[M^m(t)\right] = 1 \qquad \forall\, t \in [0,S]. \tag{11.24}$$

Now we need to prove both sides of the 'if and only if' statement:

(1) Assume $\{M^m(S)\}_{m=1}^\infty$ are uniformly integrable. By [**21**, Corollary 3.19]

$$\lim_{m\to\infty} M^m(S) = \exp\left(\sum_{i=1}^n \int_0^S \beta_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^n \int_0^S \beta_i^2(v)\,dv\right)$$

a.s. and in L^1

Taking expectations:

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^n \int_0^S \beta_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^n \int_0^S \beta_i^2(v)\,dv\right)\right] = \lim_{m\to\infty} \mathbb{E}\left[M^m(S)\right] = 1$$

by (11.24) above.

[13] A local martingale may be defined [**41**, p. 462] or [**30**, Definition 1.7] as follows: a stochastic process $X = (X(t))_{t\geq 0}$ is said to be a local martingale on (Ω, F, Q) if there exists an increasing sequence τ_n of stopping times such that τ_n tends to T a.s., and for every n the process $X^n(t)$, given by the formula $X^n(t) = X(t \wedge \tau_n)$, is a martingale.

(2) Conversely, assume

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\int_0^S \beta_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^S \beta_i^2(v)\,dv\right)\right] = 1$$

then this is a martingale and so

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\int_0^S \beta_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^S \beta_i^2(v)\,dv\right)\middle|\, F_{T_m}\right] = M^m(S)$$

Hence by [**21**, Theorem 3.20] $\{M^m(S)\}_{m=1}^{\infty}$ are uniformly integrable. $\square$

LEMMA 2.2. Assume that Conditions 1–3 hold for a fixed set $\{S_1, \ldots, S_n\} \in [0, \tau]$, $0 < S_1 < \ldots < S_n \leq \tau$. Also assume that the following four conditions of Lemma 1.1 hold:

(i) $$\begin{bmatrix} b(t,S_1) \\ \vdots \\ b(t,S_n) \end{bmatrix} + \begin{bmatrix} a_1(t,S_1) & \ldots & a_n(t,S_1) \\ & \vdots & \\ a_1(t,S_n) & \ldots & a_n(t,S_n) \end{bmatrix} \begin{bmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{a.e. } Q \times \lambda$$

(ii) $$\int_0^{S_1} \gamma_i^2(v)\,dv < +\infty \qquad \text{a.e. } Q \ \text{ for } i = 1, \ldots, n$$

(iii) $$\mathbb{E}^Q\left[\exp\left(\sum_{i=1}^{n}\int_0^{S_1} \gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1} \gamma_i^2(v)\,dv\right)\right] = 1$$

(iv) $$\mathbb{E}^Q\left[\exp\left(\sum_{i=1}^{n}\int_0^{S_1} (a_i(v,y) + \gamma_i(v))\,dz_i(v)\right.\right.$$
$$\left.\left. -\frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1} (a_i(v,y) + \gamma_i(v))^2\,dv\right)\right] = 1 \quad \text{for } y \in \{S_1, \ldots, S_n\}$$

Then $\gamma_i(t)$ for $i = 1, \ldots, n$, satisfying the above conditions are unique (up to $\lambda \times Q$ equivalence) if and only if

$$A(t) \equiv \begin{bmatrix} a_1(t,S_1) & \ldots & a_n(t,S_1) \\ & \vdots & \\ a_1(t,S_n) & \ldots & a_n(t,S_n) \end{bmatrix}$$

is singular with $(\lambda \times Q)$ measure zero.

PROOF. First, assume $A(t)$ is singular with $(\lambda \times Q)$ measure zero. Since condition (i) holds we have:

$$\begin{bmatrix} a_1(t,S_1) & \dots & a_n(t,S_1) \\ & \vdots & \\ a_1(t,S_n) & \dots & a_n(t,S_n) \end{bmatrix} \begin{bmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{bmatrix} = A(t) \begin{bmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{bmatrix} = - \begin{bmatrix} b(t,S_1) \\ \vdots \\ b(t,S_n) \end{bmatrix}$$

Singularity with measure zero implies $A(t)$ is invertible; hence we may solve for unique (up to $\lambda \times Q$) $\gamma_i(t)$s ; $i = 1, \dots, n$, as:

$$\begin{bmatrix} \gamma_1(t) \\ \vdots \\ \gamma_n(t) \end{bmatrix} = -A(t)^{-1} \begin{bmatrix} b(t,S_1) \\ \vdots \\ b(t,S_n) \end{bmatrix}$$

Conversely, define Σ as the set on which $A(t)$ is singular, that is $\Sigma \equiv \{t \times \omega \in [0,S] \times \Omega : A(t) \text{ is singular}\}$. Assume that Σ has $(\lambda \times Q)(\Sigma) > 0$. Then proof is by contradiction. Hence, having assumed $A(t)$ is singular, we wish to show that the functions satisfying (i)–(iv) are not unique.

Firstly, since (i)–(iv) are assumed to hold, we have a vector of functions $(\gamma_1(t), \dots, \gamma_n(t))$ satisfying these conditions.

Part 1 We wish to show that there exists a bounded, adapted and measurable vector of functions $(\delta_1(t), \dots, \delta_n(t))$ which is non-zero on Σ and satisfies:

$$A(t) \begin{bmatrix} \delta_1(t) \\ \vdots \\ \delta_n(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

and where $g(t)$, defined as:

$$g(t) \equiv \exp\left(\sum_{i=1}^{n} \int_0^t \delta_i(v)\, dz_i(v) - \sum_{i=1}^{n} \int_0^t \delta_i(v)\gamma_i(v)\, dv - \frac{1}{2} \sum_{i=1}^{n} \int_0^t \delta_i^2(v)\, dv \right)$$

is bounded a.e. on Q.

Let $\Sigma_i = \{(t,\omega) : A(t) \text{ has rank } i\}$. Σ_i has the following properties:

- Σ_i is a measurable set,
- $\Sigma = \bigcup_{i=1}^{n-1} \Sigma_i$,
- $\Sigma_i \cap \Sigma_j = \emptyset$ for $i \neq j$,

Fix $\eta > 0$. On each Σ_i set $\delta_i^\eta(t)$, $i = 1, \dots, n$, to be a solution to[14]:

$$A(t) \begin{bmatrix} \delta_1^\eta(t) \\ \vdots \\ \delta_n^\eta(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

[14]Here we iteratively build vector $\delta^\eta(t)$. That is for each Σ_i we determine element $\delta_i^\eta(t)$, until the entire vector has been determined.

such that $\delta_i^\eta(t)$ is bounded by $\min\{\eta, 1/\gamma_i(t)\, i = 1,\ldots,n\}$. Also, let $\delta_i^\eta(t)$ be zero on Σ^c; $i = 1,\ldots,n$. By construction, $\delta_i^\eta(t)$ are adapted, measurable and bounded by η. Also:

$$\begin{aligned}
&\left|\sum_{i=1}^n \int_0^t \delta_i^\eta(v)\gamma_i(v)\,dv + \frac{1}{2}\sum_{i=1}^n \int_0^t \delta_i^\eta(v)^2\,dv\right| \\
&\leq\ nt + \tfrac{1}{2}nt\eta^2 \\
&=\ [2+\eta^2]\tfrac{1}{2}\,nt \qquad \text{a.e. } Q
\end{aligned}$$

Set $\alpha = \inf\{j \in \{1,2,3,\ldots\} : (\frac{1}{2})^{2j} S < 1\}$ and define inductively the following stopping times:

$$\tau_1 = \inf\left(t \in [0,S] : \sum_{i=1}^n \int_0^t \delta_i^{(1/2)^\alpha}(v)\,dz_i(v) \geq \frac{1}{2}\right)$$

$$\tau_j = \inf\left(t \in [0,S] : \sum_{i=1}^n \int_{\tau_{j-1}}^t \delta_i^{(1/2)^{2j+\alpha}}(v)\,dz_i(v) \geq \left(\frac{1}{2}\right)^j\right) \qquad \text{for } j = 2,3,4,\ldots$$

We need to show that the claim $Q(\lim_{j\to\infty}\tau_j = S) = 1$ is true. First consider:

$$\begin{aligned}
Q(\tau_j < S\,|F_{\tau_{j-1}}) &\leq Q\left(\left|\sum_{i=1}^n \int_{\tau_{j-1}}^S \delta_i^{(1/2)^{2j+\alpha}}(v)\,dz_i(v)\right| \geq \left(\frac{1}{2}\right)^j \,\middle|\, F_{\tau_{j-1}}\right) \\
&\leq \frac{1}{(1/2)^{2j}} \int_{\tau_{j-1}}^S \left(\delta_i^{(1/2)^{2j+\alpha}}(v)\right)^2 dv \\
&\qquad \text{by Chebyshev's Inequality} \\
&\leq \frac{1}{(1/2)^{2j}} \int_{\tau_{j-1}}^S \left((\tfrac{1}{2})^{2j+\alpha}\right)^2 dv \\
&\qquad \text{by the boundedness of } \delta_i^\eta(v) \\
&= \frac{1}{(1/2)^{2j}} \left((\tfrac{1}{2})^{2j+\alpha}\right)^2 (S - \tau_{j-1}) \\
&\leq \frac{1}{(1/2)^{2j}} \left((\tfrac{1}{2})^{2j+\alpha}\right)^2 S \\
&= \left(\frac{1}{2}\right)^{2j}\left(\frac{1}{2}\right)^{2\alpha} S \\
&< \left(\frac{1}{2}\right)^{2j} \qquad \text{by choice of } \alpha \text{ such that } (\tfrac{1}{2})^{2\alpha} S < 1
\end{aligned}$$

and hence $\mathbb{E}\left[Q(\tau_j < S|F_{\tau_{j-1}})\right] = Q(\tau_j < S) < (\frac{1}{2})^{2j}$. Now consider[15]:

[15] This result is by Fatou's Lemma, e.g. see [**41**].

$$Q\Big(\lim_{j\to\infty}\tau_j < S\Big) < Q\bigg(\bigcap_{j=1}^{\infty}(\tau_j < S)\bigg) \le \inf\big[Q(\tau_j < S) : j = 1,2,3,\ldots\big] = 0$$

Hence:

$$Q\Big(\lim_{j\to\infty}\tau_j = S\Big) = 1 - Q\Big(\lim_{j\to\infty}\tau_j < S\Big) = 1$$

and we have proved the claim. Set

$$\delta_i(t) = \sum_{j=0}^{\infty} 1^{(t)}_{[\tau_j,\tau_{j+1}]}\, \delta_i^{(1/2)^{2j+\alpha}}(t) \qquad \text{for } i = 1,\ldots,n$$

By this construction of $\delta_i(t)$ we know it is bounded, adapted, measurable and satisfies:

$$A(t)\begin{bmatrix}\delta_1(t)\\ \vdots \\ \delta_n(t)\end{bmatrix} = \begin{bmatrix}0\\ \vdots \\ 0\end{bmatrix} \qquad \text{a.e. } \lambda\times Q$$

Also for all $t \in [0,S]$ we have:

$$\left|\sum_{i=1}^{n}\int_0^t \delta_i(v)\,dz_i(v)\right| \le \sum_{j=0}^{\infty}\Big(\frac{1}{2}\Big)^j = 2$$

and so

$$g(t) = \exp\left(\sum_{i=1}^{n}\int_0^t \delta_i(v)\,dz_i(v) - \sum_{i=1}^{n}\int_0^t \delta_i(v)\gamma_i(v)\,dv - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \delta_i^2(v)\,dv\right)$$

is bounded a.e. $\lambda\times Q$. Part 1 is complete.

Part 2 Now to complete the proof, we wish to show that $(\gamma_1(t)+\delta_1(t)$ $,\ldots,\gamma_n(t)+\delta_n(t))$ satisfies (i)–(iv).

Consider condition (i). We need to show that:

$$\begin{bmatrix}b(t,S_1)\\ \vdots \\ b(t,S_n)\end{bmatrix} + A(t)\begin{bmatrix}\gamma_1(t)+\delta_1(t)\\ \vdots \\ \gamma_n(t)+\delta_n(t)\end{bmatrix} = \begin{bmatrix}0\\ \vdots \\ 0\end{bmatrix} \qquad \text{a.e. } Q\times\lambda$$

Since $(\gamma_1(t),\ldots,\gamma_n(t)$ satisfies (i) we have:

$$\begin{bmatrix}b(t,S_1)\\ \vdots \\ b(t,S_n)\end{bmatrix} + A(t)\begin{bmatrix}\gamma_1(t)\\ \vdots \\ \gamma_n(t)\end{bmatrix} = \begin{bmatrix}0\\ \vdots \\ 0\end{bmatrix} \qquad \text{a.e. } Q\times\lambda$$

and by construction above

$$A(t)\begin{bmatrix}\delta_1(t)\\ \vdots\\ \delta_n(t)\end{bmatrix}=\begin{bmatrix}0\\ \vdots\\ 0\end{bmatrix}$$

Hence condition (i) is satisfied.

To satisfy condition (ii) we require

$$\int_0^{S_1}\Big(\gamma_i(v)+\delta_i(v)\Big)^2\,dv<+\infty\qquad \text{a.e. } Q,\ i=1,\dots,n$$

Expanding the integration, we have:

$$\begin{aligned}&\int_0^{S_1}\Big(\gamma_i(v)+\delta_i(v)\Big)^2\,dv\\ &=\int_0^{S_1}\gamma_i^2(v)\,dv+2\int_0^{S_1}\gamma_i(v)\delta_i(v)\,dv+\int_0^{S_1}\delta_i^2(v)\,dv\end{aligned}$$

Since $\gamma_i(v)$ satisfies condition (ii) and from Part 1 we know

$$\int_0^t\delta_i(v)\gamma_i(v)\,dv+\frac{1}{2}\int_0^t\delta_i^2(v)\,dv$$

is bounded, we conclude $(\gamma_i(t)+\delta_i(t))$, $i=1,\dots,n$, satisfy condition (ii).

Condition (iii) requires:

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^n\int_0^{S_1}\gamma_i(v)+\delta_i(v)\,dz_i(v)-\frac{1}{2}\sum_{i=1}^n\int_0^{S_1}\big(\gamma_i(v)+\delta_i(v)\big)^2\,dv\right)\right]=1$$

To show that this condition holds, follow the proof of Lemma 2.1. Define:

$$\begin{aligned}T_m&=\inf\left\{t\in[0,S]:\ \mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^n\int_0^t\big(\gamma_i(v)+\delta_i(v)\big)^2\,dv\right)\right]\geq m\right\}\\ M^m(t)&=\exp\left(\sum_{i=1}^n\int_0^{T_m\wedge t}(\gamma_i(v)+\delta_i(v))dz_i(v)\right.\\ &\qquad\left.-\frac{1}{2}\sum_{i=1}^n\int_0^{T_m\wedge t}(\gamma_i(v)+\delta_i(v))^2dv\right)\end{aligned}$$

and we need to show that $\{M^m(S)\}_{m=1}^\infty$ is uniformly integrable.

We have

$$\begin{aligned} M^m(S) &= \exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i^2(v)\,dv\right) \\ &\quad \times\exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i(v)dz_i(v) - \sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i(v)\delta_i(v)dv\right. \\ &\quad \left. -\frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i^2(v)dv\right) \end{aligned}$$

From Part 1 we know that

$$\exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i(v)\,dz_i(v) - \sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i(v)\delta_i(v)\,dv - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i^2(v)\,dv\right)$$

is bounded by some value, say $K > 0$. Hence:

$$0 \le M^m(S) \le K\exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i^2(v)\,dv\right)$$

and since $\gamma_i(t)$ satisfies condition (iii), the right-hand side is uniformly integrable and hence it may be shown that $M^m(S)$ is uniformly integrable.

Finally to show that condition (iv) is satisfied, we must show:

$$\mathbb{E}\left[\exp\left(\sum_{i=1}^{n}\int_0^{S_1}\big(\Phi(v,y)\big)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{S_1}\big(\Phi(v,y)\big)^2\,dv\right)\right] = 1 \qquad \text{for } y \in \{S_1,\dots,S_n\}$$

where $\Phi(v,y) = a_i(v,y) + \gamma_i(v) + \delta_i(v)$. Again, following the proof of Lemma 2.1 we define:

$$T_m = \inf\left\{t \in [0,S] : \mathbb{E}\left[\exp\left(\frac{1}{2}\sum_{i=1}^{n}\int_0^{t}\big(\Phi(v,y)\big)^2\,dv\right) \ge m\right]\right\}$$

$$M^m(t) = \exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge t}\big(\Phi(v,y)\big)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge t}\big(\Phi(v,y)\big)^2\,dv\right)$$

and the condition holds if $\{M^m(S)\}_{m=1}^{\infty}$ is uniformly integrable. $M^m(S)$ may be decomposed as:

$$
\begin{aligned}
&M^m(S)\\
&= \exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\big(a_i(v,y)+\gamma_i(v)\big)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\big(a_i(v,y)+\gamma_i(v)\big)^2\,dv\right)\\
&\quad\times\exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i(v)\,dz_i(v) - \sum_{i=1}^{n}\int_0^{T_m\wedge S}\big(a_i(v,y)\delta_i(v)+\gamma_i(v)\delta_i(v)\big)\,dv\right.\\
&\quad\left.-\frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i^2(v)\,dv\right)
\end{aligned}
$$

By choice of $\delta_i(v)$ in Part 1 we have

$$\sum_{i=1}^{n}\int_0^{T_m\wedge S} a_i(v,y)\delta_i(v)\,dv = 0$$

and

$$\exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i(v)\,dz_i(v) - \sum_{i=1}^{n}\int_0^{T_m\wedge S}\gamma_i(v)\delta_i(v)\,dv - \frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}\delta_i^2(v)\,dv\right)$$

is bounded by some value, say $K > 0$. Therefore we may write:

$$
\begin{aligned}
0 \le M^m(S) \quad \le \quad & K\exp\left(\sum_{i=1}^{n}\int_0^{T_m\wedge S}(a_i(v,y)+\gamma_i(v))dz_i(v)\right.\\
&\left.-\frac{1}{2}\sum_{i=1}^{n}\int_0^{T_m\wedge S}(a_i(v,y)+\gamma_i(v))^2 dv\right)
\end{aligned}
$$

and since $\gamma_i(t)$ satisfies condition (iv), the right-hand side is uniformly integrable and hence it may be shown that $M^m(S)$ is uniformly integrable.

□

Hence, by the proofs of Lemma 2.1 and Lemma 2.2 we have proved Proposition 2. □

Conditions 1–5 impose restrictions on the market prices of risk, $\gamma_i(t; S_1, \dots S_n),\ i = 1, \dots, n$, which results in restrictions on the drifts of the forward rate processes $\{\alpha(\cdot, S_1), \dots, \alpha(\cdot, S_n)\}$. These restrictions are required to guarantee the existence of the unique equivalent martingale probability measure for relative bond prices $\{Z(t, S_1), \dots, Z(t, S_n)\}$, $0 < S_1 < S_2 < \dots < S_n \le \tau$.

11.3.2. Model dynamics under the martingale measure. Let us now determine the dynamics of the forward rate and bond price under this equivalent martingale probability measure. In (11.23) the Brownian motions with respect to the equivalent martingale measure are defined as:

$$d\tilde{z}_i^{S_1,\,...\,,\,S_n}(t) = dz_i(t) - \gamma_i(t; S_1, \ldots, S_n)\,dt \quad i = 1, \ldots, n \tag{11.25}$$

Substituting into (11.4), the dynamics of $f(t,T)$ under the equivalent martingale measure are:

$$\begin{aligned} f(t,T) = f(0,T) &+ \int_0^t \alpha(v,T)\,dv + \sum_{i=1}^n \int_0^t \sigma_i(v,T)\gamma_i(v; S_1, \ldots, S_n)\,dv \\ &+ \sum_{i=1}^n \int_0^t \sigma_i(v,T)\,d\tilde{z}_i^{S_1,\,...\,,\,S_n}(v) \quad \text{a.e. } \widetilde{Q}_{S_1,\ldots,S_n} \;\; \forall\;\; 0 \le t \le T \end{aligned} \tag{11.26}$$

To determine the bond price process under the equivalent martingale measure, substitute (11.25) into (11.13) and make use of the market price of risk equation (11.20) to give:

$$\begin{aligned} dP(t,T) &= P(t,T)\big(r(t) + b(t,T)\big)\,dt \\ &\quad + P(t,T)\sum_{i=1}^n a_i(t,T)\left(d\tilde{z}_i^{S_1,\,...\,,\,S_n}(t) + \gamma_i(v; S_1, \ldots, S_n)\,dt\right) \\ &= P(t,T)r(t)\,dt + P(t,T)\sum_{i=1}^n a_i(t,T)\,d\tilde{z}_i^{S_1,\,...\,,\,S_n}(t) \\ &\qquad\qquad\qquad \text{a.e. } \widetilde{Q}_{S_1,\ldots,S_n} \;\; \text{for } T \in S_1, \ldots, S_n \end{aligned}$$

Also from (11.12), the logarithm bond price process, we have:

$$\begin{aligned} P(t,T) = P(0,T)\exp\Bigg(&\int_0^t r(v)\,dv - \frac{1}{2}\sum_{i=1}^n \int_0^t a_i^2(v,T)\,dv \\ &+ \sum_{i=1}^n \int_0^t a_i(v,T)\,d\tilde{z}_i^{S_1,\,...\,,\,S_n}(v)\Bigg) \\ &\qquad \text{a.e. } \widetilde{Q}_{S_1,\ldots,S_n} \;\; \text{for } T \in S_1, \ldots, S_n \end{aligned} \tag{11.27}$$

Similarly, the relative bond price under the equivalent martingale measure may be determined from (11.16) as:

$$Z(t,T) = Z(0,T)\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\int_0^t a_i^2(v,T)\,dv + \sum_{i=1}^{n}\int_0^t a_i(v,T)\,d\tilde{z}_i^{S_1,\,...,\,S_n}(v)\right) \quad \text{a.e. } \widetilde{Q}_{S_1,\dots,S_n} \text{ for } T \in S_1,\dots,S_n \tag{11.28}$$

Although these processes evolve in a risk-neutral economy, the forward rate process (11.26) (and since $r(t) = f(t,t)$, also the short-term interest rate process) displays explicit dependence on the market prices of risk, $\gamma_i(v; S_1,\dots,S_n)$, $i = 1,\dots,n$. Hence, to price any security depending on either the short or forward interest rate, the market prices of risk must be known.

11.4. Eliminating the market prices of risk

Equations (11.26)–(11.28) depend on a specific set of n bond maturities, with maturities $\{S_1,\dots,S_n\}$, for which the market prices of risk exist. These are the n bond maturities which define the equivalent martingale measure.

By making an additional assumption, specified in the condition below, we show that there exists a unique equivalent martingale measure simultaneously making relative bond prices of all maturities martingales. This allows the dependence on n specific bond maturities to be eliminated and pricing equations (11.26)–(11.28) become entirely independent of the number and maturity of bonds used to determine the equivalent martingale measure.

CONDITION 6. *Common equivalent martingale measures.* Given Conditions 1 - 3, let Conditions 4 and 5 hold for *all* bond maturities $S_1,\dots,S_n \in [0,\tau]$ with $0 < S_1 < \dots < S_n \le \tau$. Also, $\widetilde{Q} = \widetilde{Q}_{S_1,\dots,S_n}$.

PROPOSITION 3. *Uniqueness of the equivalent martingale probability measure across all bond maturities. Given:*

- *a family of forward rate drifts* $\{\alpha(\cdot,T) : T \in [0,\tau]\}$,
- *a family of forward rate volatilities* $\{\sigma_i(\cdot,T) : T \in [0,\tau]\}$, $i = 1,\dots,n$

which satisfy Conditions 1–5, then the following are equivalent:

(11.29a) $\widetilde{Q}$, *defined by* $\widetilde{Q} = \widetilde{Q}_{S_1,\ldots,S_n}$ *for any* $S_1, \ldots, S_n \in [0, \tau]$ *is the unique equivalent probability measure such that* $Z(t,T)$ *is a martingale for all* $T \in [0, \tau]$ *and* $t \in [0, T]$.

(11.29b) $\gamma_i(t; S_1, \ldots, S_n) = \gamma_i(t; T_1, \ldots, T_n)$ *for* $i = 1, \ldots, n$ *and* $S_1, \ldots, S_n, T_1, \ldots, T_n \in [0, \tau], t \in [0, \tau]$ *such that* $0 \leq t < S_1 < \ldots < S_n \leq \tau,\ 0 \leq t < T_1 < \ldots < T_n \leq \tau$.

(11.29c) $\alpha(t,T) = -\sum_{i=1}^{n} \sigma_i(t,T)\big(\phi_i(t) - \int_t^T \sigma_i(t,v)\,dv\big)\ \forall\, T \in [0,\tau]$, $t \in [0, T]$ *where for* $i = 1, \ldots, n$ $\ \phi_i(t) = \gamma_i(t; S_1, \ldots, S_n)$ *for any* $S_1, \ldots, S_n \in [0, \tau]$ *and* $t \in [0, S_1]$.

Hence the existence of the unique equivalent probability measure $\widetilde{Q}$ which makes all relative bond prices martingales, is equivalent to the condition making market prices of risk independent of the specific bond maturities chosen, which is equivalent to the above restriction on the drift of the forward rate process. All of the above conditions ensure an arbitrage-free framework.

Examine the derivation and resulting implications of each of the above conditions in turn. (11.29a) implies the martingale approach to bond pricing. $Z(t,T)$ is a martingale with respect to the equivalent probability measure $\widetilde{Q}$, hence:

$$Z(t,T) = \mathbb{E}^{\widetilde{Q}}\left[Z(T,T)|F_t\right]$$
$$\Rightarrow \frac{P(t,T)}{B(t)} = \mathbb{E}^{\widetilde{Q}}\left[\left.\frac{P(T,T)}{B(T)}\right| F_t\right] = \mathbb{E}^{\widetilde{Q}}\left[\left.\frac{1}{B(T)}\right| F_t\right]$$

Then, by Girsanov's Theorem:

$$P(t,T) = B(t)\ \mathbb{E}^{Q}\left[\left.\frac{d\widetilde{Q}}{dQ}\frac{1}{B(T)}\right| F_t\right] \tag{11.30}$$

where $\frac{d\widetilde{Q}}{dQ}$ is the Radon–Nikodym derivative defining the equivalent martingale probability measure. From (11.22) and the above proposition, showing the independence of the market prices of risk of bond maturity T and equivalently the uniqueness of the equivalent probability measure across all bond maturities, the Radon–Nikodym derivative may be written as:

$$\frac{d\widetilde{Q}}{dQ} = \exp\left(\sum_{i=1}^{n}\int_0^T \phi_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^{n}\int_0^T \phi_i^2(v)\,dv\right)$$

and so (11.30) becomes:

(11.31)
$$P(t,T) = B(t)\ \mathbb{E}^Q\left[\left.\frac{\exp\left(\sum_{i=1}^n \int_0^T \phi_i(v)\,dz_i(v) - \frac{1}{2}\sum_{i=1}^n \int_0^T \phi_i^2(v)\,dv\right)}{B(T)}\right| F_t\right]$$

When pricing under the original market measure Q, the bond price is explicitly dependent on the money market account $B(T)$ and the market prices of risk $\phi_i(t),\ i = 1, \ldots, n$. This introduces an implicit dependence on:

- forward rate drifts under the market measure $\{\alpha(\cdot, T) : T \in [0, \tau]\}$,
- forward rate volatilities $\{\sigma_i(\cdot, T) : T \in [0, \tau]\},\ i = 1, \ldots, n$,
- initial forward rate curve $\{f(0, T) : T \in [0, \tau]\}$.

Condition (11.29b), requiring the independence of the market prices of risk of bond maturity, is a necessary condition for the absence of arbitrage. It is a standard condition, imposed by many earlier models (e.g. Vasicek [**50**], Cox, Ingersoll and Ross [**18**] and Brennan and Schwartz [**10**]) to derive the fundamental partial differential equation for contingent claim valuation.

Condition (11.29c) imposes a restriction on the functional form of the family of drift processes $\{\alpha(\cdot, T) : T \in [0, \tau]\}$, which is required to ensure the existence of the unique equivalent martingale probability measure. Not all possible forward rate drift processes will comply with this condition.

Examine closely the derivation of (11.29c). By (11.29b) the market prices of risk are independent of the set of bond maturities specified, so (11.20) is equivalent to (11.17) with $\gamma_i(t) = \gamma_i(t; S_1, \ldots, S_n) = \phi_i(t)$. Making use of the definitions of $a_i(t, T)$ and $b(t, T)$ in (11.11), equation (11.20) becomes:

$$\begin{aligned}
b(t,T) &= \sum_{i=1}^n a_i(t,T)\,(-\phi_i(t)) \\
\Rightarrow \int_t^T \alpha(t,v)\,dv &= \frac{1}{2}\sum_{i=1}^n a_i^2(t,T) + \sum_{i=1}^n a_i(t,T)\phi_i(t)
\end{aligned}$$

Differentiating with respect to T yields the required form of the forward rate drift restriction:

$$\begin{aligned}
\alpha(t,T) &= -\sum_{i=1}^n a_i(t,T)\sigma_i(t,T) - \sum_{i=1}^n \sigma_i(t,T)\phi_i(t) \\
\Rightarrow \alpha(t,T) &= -\sum_{i=1}^n \sigma_i(t,T)\big(\phi_i(t) + a_i(t,T)\big) \\
&= -\sum_{i=1}^n \sigma_i(t,T)\Big(\phi_i(t) - \int_t^T \sigma_i(t,v)\,dv\Big)
\end{aligned}$$

To eliminate the market prices of risk from the forward rate process (11.26), make use of this forward rate drift restriction. Integrating the restriction in (11.29c) over $[0,t]$ yields:

$$\int_0^t \alpha(v,T)\,dv = -\sum_{i=1}^n \int_0^t \sigma_i(v,T)\phi_i(v)\,dv + \sum_{i=1}^n \int_0^t \sigma_i(v,T)\int_v^T \sigma_i(v,y)\,dy\,dv$$

Substituting into (11.26) yields the forward rate process under the equivalent martingale measure in and independent of market prices of risk as:

$$\begin{aligned} f(t,T) = f(0,T) &+ \sum_{i=1}^n \int_0^t \sigma_i(v,T)\int_v^T \sigma_i(v,y)\,dy\,dv \\ &+ \sum_{i=1}^n \int_0^t \sigma_i(v,T)\,d\tilde{z}_i(v) \qquad \text{a.e. } \widetilde{Q} \quad \forall\ 0 \le t \le T \end{aligned} \tag{11.32}$$

From (11.3), we have $r(t) = f(t,t)$, and the short-term interest rate process may be expressed as:

$$\begin{aligned} r(t) = f(0,t) &+ \sum_{i=1}^n \int_0^t \sigma_i(v,t)\int_v^t \sigma_i(v,y)\,dy\,dv \\ &+ \sum_{i=1}^n \int_0^t \sigma_i(v,t)\,d\tilde{z}_i(v) \qquad \text{a.e. } \widetilde{Q} \quad \forall\ t \in [0,\tau] \end{aligned} \tag{11.33}$$

Here, the market prices of risk have been replaced by a series of forward rate volatilities of various maturities. Hence the short-term interest rate for time t is determined using all possible volatility information contained in the term structure over the time interval $[0,t]$.

The bond price in (11.27) is not an explicit function of the market prices of risk, these enter only via the short-term interest rate process $r(t)$. Hence the formulae for the bond and relative bond prices (11.27) and (11.28) remain unchanged once the market prices of risk have been eliminated, except that the Brownian motion no longer depends on the specific n bond maturities chosen, that is $d\tilde{z}_i^{S_1,\dots,S_n}(\cdot) \equiv d\tilde{z}_i(\cdot)$. Additionally the formulae may be applied to bonds of all maturities T, $T \in [0,\tau]$.

The original formulation of the CIR [**18**] model (see Chapter 2) begins with a characterisation of an equilibrium economy. The functional form of the short-term interest rate process and the market price of risk are determined from within this economy. CIR criticise arbitrage pricing theory on the grounds that it exogenously specifies the functional form of the short-term interest rate and market prices of risk, independently of an underlying equilibrium economy. They show that this may lead to inconsistencies and a model admitting arbitrage.

However, in this model the interdependence of the short-term interest rate process, bond price process and market prices of risk is easily seen in equations (11.26)–(11.28). Additionally, HJM make use of information contained in the bond price process to eliminate the market prices of risk from the pricing formulae. This makes their general pricing framework immune to the criticism of CIR.

11.5. The problem with forward rates

By (11.32) we see that the forward rate process is completely specified by the volatility functions $\sigma_i(\cdot,\cdot)$, $i = 1, \ldots, n$. Consider a framework with one source of uncertainty and so one volatility parameter. For practical implementation it is desirable to apply a lognormal volatility structure for all forward rates [**45**]. This is because market prices of caps and swaptions assume a lognormal structure of forward rates. Hence set $\sigma_1(t,T) = \sigma f(t,T)$, where $\sigma > 0$ is a constant. However, under this volatility structure (11.32) becomes:

$$df(t,T) = \sigma^2 f(t,T) \int_t^T f(t,y)\, dy\, dt + \sigma f(t,T)\, d\tilde{z}(t) \tag{11.34}$$

Here, the drift of the forward rate grows as the square of the forward rate [**49**] and causes the forward rate to explode in finite time. Therefore for calibration purposes an upper bound needs to be imposed:

$$\sigma_1(t,T) = \sigma f(t,T) \min\{M, f(t,T)\}$$

This problem is not particular to the HJM model, but rather a characteristic of all lognormal models of instantaneous forward rates.

11.6. Unifying framework for contingent claim valuation

HJM impose conditions on the forward rate process to ensure that a unique, equivalent martingale probability measure exists and hence the process is consistent with an arbitrage-free market . This implies that the market is complete and contingent claims may be valued using an approach detailed by Harrison and Kreps [**23**] and Harrison and Pliska [**24**]. Harrison and Pliska examine martingale theory within a continuous trading environment, presenting a general methodology for contingent claim valuation. First consider some important concepts and definitions characterising this methodology:

(1) If $\mathbb{Q}$ is a set of probability measures, equivalent to initial probability measure Q and making discounted prices martingales, then by Harrison and Pliska [**24**, Corollary 3.36]: if $\mathbb{Q}$ is a singleton[16] then the market is complete.

[16] A singleton is a set having a single element.

(2) In a complete market there are enough non-redundant securities being traded [**22**], such that every integrable contingent claim is attainable.
- By attainable, we mean that there exists some trading strategy, requiring an initial investment and thereafter producing the same cash flows as the contingent claim.
- A trading strategy may be viewed as some portfolio of securities, where the number of units of each security held changes through time.

(3) Define a contingent claim as a random variable $X : \Omega \to R,\ X \geq 0$ which is F_{T_1} measurable[17]. Since contingent claim X must be integrable, we require $\mathbb{E}\left[\frac{X}{B(T_1)}\right] < +\infty$.

(4) Denote an admissible, self-financing trading strategy Υ by $\{N_0(t), N_1(t), \ldots, N_n(t)\}$ where $N_i(t)$ is the quantity of asset P_i, $i = 1, \cdots, n$ in the portfolio at time t.
- A trading strategy is admissible if it is self-financing and the value of the associated portfolio remains non negative through time. This implies that an investor enters the trading strategy with positive wealth and is never in a position of debt.
- A self-financing trading strategy is one where changes in value of the portfolio are due to capital gains (changes in value of the instruments held) only, not due to cash inflows and outflows. Let $V_t(\Upsilon)$ be the time t value of trading strategy Υ, then by [**22**, Definition 1] a self-financing trading strategy is one where:

$$dV_t(\Upsilon) = \sum_{i=1}^{n} N_i(t)\, dP_i(t)$$

(5) An arbitrage opportunity is some trading strategy Υ such that $V_0(\Upsilon) = 0$ and $\mathbb{E}\left[V_{T_1}(\Upsilon)\right] > 0$. The existence of arbitrage opportunities is inconsistent with an equilibrium in the economy.

(6) Harrison and Pliska present a theorem showing that a market is free of arbitrage opportunities if and only if $\mathbb{Q}$ is not empty. Hence the absence of arbitrage opportunities is equivalent to the existence of an equivalent martingale probability measure.

(7) By Harrison and Pliska [**24**, Proposition 2.9], if X is an attainable contingent claim generated by trading strategy Υ, with time t value $V_t(\Upsilon)$ and $\widetilde{Q} \in \mathbb{Q}$ then:

[17] It is a security entitling the holder to a payment at time T_1. The magnitude of this payment depends on the history of price movements up to time T_1, hence it is F_{T_1} measurable.

$$\frac{V_t(\Upsilon)}{B(t)} = \mathbb{E}^{\widetilde{Q}}\left[\left.\frac{X}{B(T_1)}\right| F_t\right], \quad t \in [0, T_1]$$

Now, consider the economic framework characterised in the previous sections: Let Conditions 1–6 hold. By Proposition 3, there is a unique equivalent measure $\widetilde{Q}$ making all $Z(t,T)$, $T \in [0,\tau]$, $t \in [0,T]$ martingales. Since $\widetilde{Q}$ is unique, the market is complete and there exists an admissible, self-financing trading strategy, as denoted above, such that the portfolio value satisfies:

$$N_0(T_1)B(T_1) + \sum_{i=1}^{n} N_{T_i}(T_1)P(T_1, T_i) = X \quad \text{a.e. } \widetilde{Q} \tag{11.35}$$

where

$N_0(T_1)$	–	amount held in the money market account at time T_1,
$N_{T_i}(T_1)$	–	amount of bond with maturity time T_i in the self-financing strategy at time T_1,
$P(T_1, T_i)$	–	time T_1 value of a bond maturing at time T_i,
X	–	time T_1 payout of a contingent claim.

Now, since we are in a complete market where all contingent claims are attainable and where a unique equivalent martingale measure exists, we conclude that arbitrage opportunities do not exist and the time t price of a contingent claim paying X at time T_1 is given by:

$$\mathbb{E}^{\widetilde{Q}}\left[\left.\frac{X}{B(T_1)}\right| F_t\right] B(t) \tag{11.36}$$

Hence given (11.35), the trading strategy which generates X, the time t value of the portfolio is:

$$\begin{aligned}
&\mathbb{E}^{\widetilde{Q}}\left[\left.\frac{X}{B(T_1)}\right| F_t\right] B(t) \\
&= \mathbb{E}^{\widetilde{Q}}\left[\left. N_0(T_1) + \frac{N_{T_1}(T_1)}{B(T_1)} + \sum_{i=2}^{n} N_{T_i}(T_1)\frac{P(T_1,T_i)}{B(T_1)}\right| F_t\right] B(t) \\
&= \mathbb{E}^{\widetilde{Q}}\left[\left. N_0(T_1) + \frac{N_{T_1}(T_1)}{B(T_1)} + \sum_{i=2}^{n} N_{T_i}(T_1)Z(T_1,T_i)\right| F_t\right] B(t) \qquad (11.37)
\end{aligned}$$

Therefore, to value the contingent claim, the dynamics of the short-term interest rate $r(t)$ and the relative bond price $Z(t,T)$ must be known under the equivalent martingale measure. Since the market is complete, every contingent claim may be replicated by means of some admissible, self-financing strategy consisting of only the money market account and some n bonds with maturities $T_1, \ldots, T_n \in [0, \tau]$. From this we conclude that all contingent claims may be valued.

11.7. Ho and Lee model within the HJM framework

By way of an illustrative example, consider the continuous time equivalent of the Ho and Lee [**27**] model (studied in Chapter 10) within the above specified framework.

11.7.1. The model specifications. Consider a model with a single Brownian motion and hence single volatility parameter $\sigma_1(\omega, t, T) \equiv \sigma > 0$ where σ is a constant. Assume Conditions 1–6 are satisfied with $\{f(0,T) : T \in [0,\tau]\}$ the initial forward rate curve and $\phi(\omega, t),\ t \in [0,\tau]$ the market price of risk corresponding to the single source of randomness. From (11.32) the forward rate process with respect to the equivalent martingale measure is:

$$\begin{aligned} f(t,T) &= f(0,T) + \int_0^t \sigma^2 (T - v)\, dv + \int_0^t \sigma\, d\tilde{z}(v) \\ &= f(0,T) + \left[\sigma^2 T v - \tfrac{1}{2}\sigma^2 v^2\right]_{v=0}^{v=t} + \sigma\tilde{z}(t) - \sigma\tilde{z}(0) \\ &= f(0,T) + \sigma^2 t\big(T - \tfrac{1}{2}t\big) + \sigma\tilde{z}(t) \end{aligned} \tag{11.38}$$

and hence the short-term interest rate process is specified as:

$$r(t) \equiv f(t,t) = f(0,t) + \frac{\sigma^2 t^2}{2} + \sigma\tilde{z}(t) \tag{11.39}$$

Since there are no restrictions placed on the evolution of the short-term or forward interest rates, there exists a positive probability of these rates being negative. To determine the bond price dynamics, substitute (11.38) into (11.2) to give:

$$P(t,T) = \exp\left(-\int_t^T \Big(f(0,s) + \sigma^2 t\Big(s - \tfrac{1}{2}t\Big) + \sigma\tilde{z}(t)\Big)\, ds\right)$$

From the definition of the forward rate (11.1), we have:

$$\begin{aligned} f(0,s) &= -\frac{\partial \ln P(0,s)}{\partial s} \\ \Rightarrow -\int_t^T f(0,s)\, ds &= \int_t^T \frac{\partial \ln P(0,s)}{\partial s}\, ds \\ &= \ln\frac{P(0,T)}{P(0,t)} \end{aligned}$$

and so:

$$\begin{aligned} P(t,T) &= \exp\left(\ln\frac{P(0,T)}{P(0,t)} - \left[\tfrac{1}{2}\,\sigma^2 t s^2 - \tfrac{1}{2}\,\sigma^2 t^2 s\right]_{s=t}^{s=T} - \sigma(T-t)\tilde{z}(t)\right) \\ &= \frac{P(0,T)}{P(0,t)}\,\exp\left(-\tfrac{1}{2}\,\sigma^2 t T(T-t) - \sigma(T-t)\tilde{z}(t)\right) \qquad (11.40) \end{aligned}$$

11.7.2. Pricing contingent claims. Define the following notation:

$C(t)$ – time t value of a European call option on bond $P(t,T)$,
K – option exercise price,
t^* – option expiry date, $0 \le t \le t^* \le T$.

At expiry of the option, its value is:

$$C(t^*) = \max\left\{P(t^*,T) - K, 0\right\}$$

Making use of (11.36), the time t value of the contingent claim is[18]:

$$C(t) = \widetilde{\mathbb{E}}\left[\frac{\max\left\{P(t^*,T) - K, 0\right\}B(t)}{B(t^*)}\middle| F_t\right] \qquad (11.41)$$

An analysis, similar to that used in the initial formulation of the Black–Scholes valuation formula, yields the contingent claim price to be[19]:

$$C(t) = P(t,T)N(h_1) - KP(t,t^*)N(h_2) \qquad (11.42)$$

with

$$h_1 = \frac{\ln\left(\frac{P(t,T)}{KP(t,t^*)}\right) + \frac{1}{2}\,\sigma^2(T-t^*)^2(t^*-t)}{\sigma(T-t^*)\sqrt{t^*-t}} \qquad (11.43)$$

and

$$h_2 = h_1 - \sigma(T-t^*)\sqrt{t^*-t}$$

The above formula is a modification of the Black–Scholes option pricing formula where the required volatility is the standard deviation of instantaneous returns on the forward bond price, that is the standard deviation of returns at time t^* of a bond maturing at time T. This may be determined from (11.13), the bond price process within the current framework:

[18]Here $\widetilde{\mathbb{E}}[\cdot]$ denotes the expectation taken with respect to the equivalent martingale measure.

[19]Here $N(\cdot)$ is the cumulative normal distribution.

$$\begin{aligned} dP(t,T) &= [\dots]\,dt - P(t,T)\int_t^T \sigma\,dv\,dz(t) \\ &= [\dots]\,dt - P(t,T)\sigma(T-t)\,dz(t) \end{aligned}$$

and so the standard deviation of the forward bond price $\frac{P(t,T)}{P(t,t^*)} \equiv P(t^*,T)$, is $\sigma(T-t) - \sigma(t^*-t) = \sigma(T-t^*)$.

11.8. Comparison of equilibrium and arbitrage pricing

The CIR model [**18**] (studied in Chapter 2) is an example of equilibrium pricing methodology, while the HJM model makes use of arbitrage pricing. In order to compare the two methodologies, we examine the CIR model within the HJM framework.

CIR use a square root process to model the single underlying state variable, the short-term interest rate $r(t),\ t \in [0,\tau]$. Hence the short-term interest rate process is specified as:

$$dr(t) = \kappa\big(\theta(t) - r(t)\big)\,dt + \sigma\sqrt{r(t)}\,dz(t) \tag{11.44}$$

where the constants $r(0)$, κ, $\sigma > 0$, $\theta(t) > 0$ are continuous and $\{z(t) : t \in [0,\tau]\}$ is a standard Wiener process. To ensure that negative interest rates are precluded, the restriction $2\kappa\theta(t) \geq \sigma^2$ for all $t \in [0,\tau]$ is imposed, making $r(t) = 0$ an inaccessible boundary.

The time t price of a T maturity bond ($T \in [0,\tau]$, $t \in [0,T]$) is assumed to be of the form[20]:

$$P(t,T) = A(t,T)\,e^{-\bar{B}(t,T)r(t)} \tag{11.45}$$

and so, in equilibrium the bond price dynamics are represented by:

$$dP(t,T) = P(t,T)r(t)\big(1 - \lambda\bar{B}(t,T)\big)dt - P(t,T)\bar{B}(t,T)\sigma\sqrt{r(t)}\,dz(t)$$

[20]Here, $\bar{B}(\cdot,\cdot)$ is used rather than the more conventional $B(\cdot,\cdot)$, so as not to cause confusion with the money market account $B(\cdot)$, used in the derivation of the HJM model.

where[21]

$$\bar{B}(t,T) = \frac{2\big(e^{\gamma(T-t)} - 1\big)}{\big(\gamma + \kappa + \lambda\big)\big(e^{\gamma(T-t)} - 1\big) + 2\gamma} \tag{11.46a}$$

$$\ln A(t,T) = -\kappa \int_t^T \theta(s)\bar{B}(s,T)\,ds \tag{11.46b}$$

$$\gamma^2 = \big(\kappa + \lambda\big)^2 + 2\sigma^2 \tag{11.46c}$$

and λ is a constant related to the market price of risk such that $\phi(t) = -\frac{\lambda\sqrt{r(t)}}{\sigma}$. This functional form of the market price of risk is not arbitrarily specified, but is determined directly by the underlying equilibrium economy.

From (11.1), the forward rate is defined as:

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}$$

So, making use of (11.45) and (11.46b) the form of the forward rate curve is:

$$\begin{aligned} f(t,T) &= -\frac{\partial}{\partial T}\Big(\ln A(t,T) - \bar{B}(t,T)r(t)\Big) \\ &= r(t)\frac{\partial \bar{B}(t,T)}{\partial T} - \frac{\partial}{\partial T}\Big(-\kappa\int_t^T \theta(s)\bar{B}(s,T)\,ds\Big) \\ &= r(t)\frac{\partial \bar{B}(t,T)}{\partial T} + \kappa\int_t^T \theta(s)\frac{\partial \bar{B}(s,T)}{\partial T}\,ds \end{aligned} \tag{11.47}$$

While HJM allow the initial ($t = 0$) forward rate curve to be an exogenous input, here the initial forward rate curve has a pre-specified functional form. The functional form is dependent on the parameters of the specific model (i.e. κ, σ, λ and $\theta(\cdot)$) and may be determined by evaluating (11.47) at time $t = 0$:

$$f(0,T) = r(0)\frac{\partial \bar{B}(0,T)}{\partial T} + \kappa\int_0^T \theta(s)\frac{\partial \bar{B}(s,T)}{\partial T}\,ds \tag{11.48}$$

To allow any specific initial forward rate curve to be matched, CIR suggest solving (11.48) for $\{\theta(t) : t \in [0,\tau]\}$, thereby determining the time-dependent short-term interest rate drift parameters from the initial forward rate curve. However, CIR do not attempt an implementation of such a methodology or prove the existence of such a solution. Additionally, the CIR model is not consistent with all possible forward rate curves, since it imposes the restriction $2\kappa\theta(t) \geq \sigma^2$ for all $t \in [0,\tau]$. Consider this restriction in terms of the

[21] These are the functional forms of $\bar{B}(t,T)$ and $A(t,T)$ as presented in equations (2.43) and (2.44) of Chapter 2.

parameters of the forward rate process. Substituting the restriction $\theta(t) \geq \frac{\sigma^2}{2\kappa}$ into the initial forward rate process (11.48) we have[22]:

$$
\begin{aligned}
f(0,T) &= r(0)\frac{\partial \bar{B}(0,T)}{\partial T} + \kappa \int_0^T \theta(s)\frac{\partial \bar{B}(s,T)}{\partial T}\,ds \\
&\geq r(0)\frac{\partial \bar{B}(0,T)}{\partial T} + \frac{\sigma^2}{2}\int_0^T \frac{\partial \bar{B}(s,T)}{\partial T}\,ds \\
&= r(0)\frac{\partial \bar{B}(0,T)}{\partial T} + \frac{\sigma^2}{2}\frac{4\gamma}{(\gamma+\kappa+\lambda)} \\
&\quad \times\left(\frac{1}{2\gamma} - \frac{1}{(\gamma+\kappa+\lambda)(e^{\gamma T}-1)+2\gamma}\right) \\
&= r(0)\frac{\partial \bar{B}(0,T)}{\partial T} + \sigma^2\left(\frac{(e^{\gamma T}-1)}{(\gamma+\kappa+\lambda)(e^{\gamma T}-1)+2\gamma}\right) \\
&= r(0)\frac{\partial \bar{B}(0,T)}{\partial T} + \sigma^2\frac{\bar{B}(0,T)}{2}
\end{aligned}
$$

Hence only initial forward rate curves satisfying this functional form are admissible.

One of the key features of the HJM framework is the restriction placed on the forward rate drift, which ensures that profitable arbitrage is precluded. We need to derive the forward rate process in such a form so as to verify that this restriction is satisfied within the CIR equilibrium framework. Apply Ito's Lemma to (11.47), treating the forward rate as a function of the stochastic variable $r(t)$ and time t, that is $f(t,T) \equiv f(r(t),t,T)$:

[22]By definition of $\bar{B}(t,T)$ we have:

$$
\frac{\partial \bar{B}(t,T)}{\partial T} = \frac{4\gamma^2 e^{\gamma(T-t)}}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^2}
$$

and so:

$$
\begin{aligned}
\int_0^T \frac{\partial \bar{B}(s,T)}{\partial T}\,ds &= \int_0^T \frac{4\gamma^2 e^{\gamma(T-s)}}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-s)}-1)+2\gamma\right]^2}\,ds \\
&= \frac{4\gamma}{(\gamma+\kappa+\lambda)}\left.\frac{1}{(\gamma+\kappa+\lambda)(e^{\gamma(T-s)}-1)+2\gamma}\right|_{s=0}^{s=T} \\
&= \frac{4\gamma}{(\gamma+\kappa+\lambda)}\left(\frac{1}{2\gamma} - \frac{1}{(\gamma+\kappa+\lambda)(e^{\gamma T}-1)+2\gamma}\right)
\end{aligned}
$$

$$\begin{aligned}
df(t,T) &= \frac{\partial f(t,T)}{\partial r(t)}\,dr(t) + \frac{1}{2}\frac{\partial^2 f(t,T)}{\partial r(t)^2}\,dr(t)\,dr(t) + \frac{\partial f(t,T)}{\partial t}\,dt \\
&= \frac{\partial \bar{B}(t,T)}{\partial T}\,dr(t) + r(t)\frac{\partial^2 \bar{B}(t,T)}{\partial T \partial t}\,dt - \kappa\theta(t)\frac{\partial \bar{B}(t,T)}{\partial T}\,dt \\
&= r(t)\left(\frac{\partial^2 \bar{B}(t,T)}{\partial T \partial t} - \kappa\frac{\partial \bar{B}(t,T)}{\partial T}\right)dt + \sigma\sqrt{r(t)}\,\frac{\partial \bar{B}(t,T)}{\partial T}\,dz(t) \qquad (11.49)
\end{aligned}$$

This is a representation of the CIR model within the HJM framework. The short-term interest rate $r(t)$, written in terms of the forward rate, which is taken as the starting point of the HJM framework, is (from (11.47)):

$$r(t) = \left(f(t,T) - \kappa\int_t^T \theta(s)\frac{\partial \bar{B}(s,T)}{\partial T}\,ds\right)\bigg/\frac{\partial \bar{B}(t,T)}{\partial T},$$

The initial forward rate curve $\{f(0,T) : T \in [0,\tau]\}$ is an exogenous input and is used to determine $\theta(t)$, $t \in [0,\tau]$ as a solution to (11.48).

The HJM no arbitrage condition (11.29c) requires that the forward rate drift $\alpha(t,T)$ has the form:

$$\alpha(t,T) = -\sigma(t,T)\left(\phi(t) - \int_t^T \sigma(t,v)\,dv\right)$$

where $\sigma(t,T)$ is the volatility of the forward rate and $\phi(t)$ the market price of risk. Within the CIR model the market price of risk has the fixed functional form $\phi(t) = -\frac{\lambda\sqrt{r(t)}}{\sigma}$ and from (11.49) the forward rate volatility is $\frac{\partial \bar{B}(t,T)}{\partial T}\,\sigma\sqrt{r(t)}$, hence the forward rate drift must have the form:

$$\begin{aligned}
\alpha(t,T) &= -\frac{\partial \bar{B}(t,T)}{\partial T}\,\sigma\sqrt{r(t)}\left(-\frac{\lambda\sqrt{r(t)}}{\sigma} - \int_t^T \frac{\partial \bar{B}(t,v)}{\partial v}\,\sigma\sqrt{r(t)}\,dv\right) \\
&= r(t)\left(\lambda\frac{\partial \bar{B}(t,T)}{\partial T} + \sigma^2\bar{B}(t,T)\frac{\partial \bar{B}(t,T)}{\partial T}\right)
\end{aligned}$$

From (11.46a), the definition of $\bar{B}(t,T)$, we have:

$$\lambda\frac{\partial \bar{B}(t,T)}{\partial T} = \frac{4\lambda\gamma^2 e^{\gamma(T-t)}}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^2}$$

and

$$\sigma^2\bar{B}(t,T)\frac{\partial \bar{B}(t,T)}{\partial T} = \frac{8\sigma^2\gamma^2 e^{\gamma(T-t)}(e^{\gamma(T-t)}-1)}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}$$

Hence:

$$\begin{aligned}
&\lambda\frac{\partial\bar{B}(t,T)}{\partial T}+\sigma^2\bar{B}(t,T)\frac{\partial\bar{B}(t,T)}{\partial T}\\
&\quad=\frac{4\gamma^2e^{\gamma(T-t)}\left[(e^{\gamma(T-t)}-1)(2\sigma^2+\lambda^2+\lambda\kappa+\lambda\gamma)+2\lambda\gamma\right]}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}
\end{aligned}\tag{11.50}$$

However, in (11.49) the forward rate drift is represented as:

$$r(t)\left(\frac{\partial^2\bar{B}(t,T)}{\partial T\partial t}-\kappa\frac{\partial\bar{B}(t,T)}{\partial T}\right)$$

Again from (11.46a), the definition of $\bar{B}(t,T)$, we have:

$$\begin{aligned}
&\frac{\partial^2\bar{B}(t,T)}{\partial T\partial t}\\
&=\frac{\partial}{\partial t}\left[\frac{4\gamma^2e^{\gamma(T-t)}}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^2}\right]\\
&=\frac{-4\gamma^3e^{\gamma(T-t)}\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]+8\gamma^3e^{2\gamma(T-t)}(\gamma+\kappa+\lambda)}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}
\end{aligned}$$

and so:

$$\begin{aligned}
&\frac{\partial^2\bar{B}(t,T)}{\partial T\partial t}-\kappa\frac{\partial\bar{B}(t,T)}{\partial T}\\
&=\frac{-4\gamma^3e^{\gamma(T-t)}\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]+8\gamma^3e^{2\gamma(T-t)}(\gamma+\kappa+\lambda)}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}\\
&\quad-\frac{4\kappa\gamma^2e^{\gamma(T-t)}\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}\\
&=\frac{4\gamma^2e^{\gamma(T-t)}\left[(e^{\gamma(T-t)}-1)(\gamma^2-\kappa^2-\lambda\kappa+\lambda\gamma)+2\lambda\gamma\right]}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}
\end{aligned}$$

From (11.46c) we have:

$$\begin{aligned}
\gamma^2&=\kappa^2+2\lambda\kappa+\lambda^2+2\sigma^2\\
\Rightarrow\gamma^2-\kappa^2-\lambda\kappa&=2\sigma^2+\lambda^2+\lambda\kappa
\end{aligned}$$

and so

$$\begin{aligned}
&\frac{\partial^2\bar{B}(t,T)}{\partial T\partial t}-\kappa\frac{\partial\bar{B}(t,T)}{\partial T}\\
&\quad=\frac{4\gamma^2e^{\gamma(T-t)}\left[(e^{\gamma(T-t)}-1)(2\sigma^2+\lambda^2+\lambda\kappa+\lambda\gamma)+2\lambda\gamma\right]}{\left[(\gamma+\kappa+\lambda)(e^{\gamma(T-t)}-1)+2\gamma\right]^3}
\end{aligned}\tag{11.51}$$

Comparing (11.50) and (11.51) we see that:

$$\lambda \frac{\partial \bar{B}(t,T)}{\partial T} + \sigma^2 \bar{B}(t,T) \frac{\partial \bar{B}(t,T)}{\partial T} = \frac{\partial^2 \bar{B}(t,T)}{\partial T \partial t} - \kappa \frac{\partial \bar{B}(t,T)}{\partial T}$$

and the CIR forward rate drift satisfies the HJM no arbitrage condition (11.29c).

The fundamental difference between the above two approaches is that CIR fix the functional form of the market price of risk (this market price of risk is determined such that an underlying economy is in equilibrium) and derive the forward rate process endogenously. HJM take some form of the forward rate process (determined by an unrelated exogenous methodology) and use it to determine the prices of contingent claims, ensuring that profitable arbitrage is precluded.

11.9. Markovian HJM model

In term structure models such as the Vasicek [**50**] and CIR [**18**] models, the starting point is the dynamics of the short-term interest rate. The drift and volatility are specified such that the short-term interest rate is Markovian. In the HJM framework the forward rate volatility and initial forward rate curve are used to characterise the term structure. Such a specification may give rise to non-Markovian short-term interest rate dynamics. In fact many HJM-based models cannot occur in a framework of Markovian short-term interest rates. The Markovian property of the short-term interest rate is desirable since it allows for simpler numerical valuation procedures of the term structure and interest rate contingent claims since:

- The term structure at time t is a function of t, maturity T and the time t short-term interest rate.
- The evolution of the short-term interest rate may be modelled using a *recombining* tree or lattice, which has significant implications for computational efficiency.

Carverhill [**14**] and Jeffrey [**32**] characterise restrictions on the volatility structure of forward rates that lead to Markovian short-term interest rate dynamics.

11.9.1. Deterministic bond price volatility.

Carverhill requires the bond price volatility structure to be deterministic, that is $a_i(t,T) = -\int_t^T \sigma_i(t,v)\,dv,\ i = 1,\ldots,n$ are functions of t and T only. Also each $a_i(t,T)$ is assumed to be twice continuously differentiable with respect to the maturity time T.

Consider (11.33), the short-term interest rate process within the HJM framework:

$$r(t) = f(0,t) + \sum_{i=1}^{n} \int_0^t \sigma_i(v,t) \int_v^t \sigma_i(v,y)\, dy\, dv + \sum_{i=1}^{n} \int_0^t \sigma_i(v,t)\, d\tilde{z}_i(v) \tag{11.52}$$

Taking the derivative with respect to t, the dynamics of the short-term interest rate may be found as:

$$\begin{aligned}
dr(t) &= \frac{\partial f(0,t)}{\partial t}\, dt + \sum_{i=1}^{n} \sigma_i(t,t) \int_t^t \sigma_i(t,y)\, dy\, dt \\
&\quad + \sum_{i=1}^{n} \int_0^t \left(\sigma_i(v,t)\sigma_i(v,t) + \frac{\partial \sigma_i(v,t)}{\partial t} \int_v^t \sigma_i(v,y)\, dy \right) dv\, dt \\
&\quad + \sum_{i=1}^{n} \int_0^t \frac{\partial \sigma_i(v,t)}{\partial t}\, d\tilde{z}_i(v)\, dt + \sum_{i=1}^{n} \sigma_i(t,t)\, d\tilde{z}_i(t) \\
&= \left(\frac{\partial f(0,t)}{\partial t} + \sum_{i=1}^{n} \int_0^t \left(\sigma_i^2(v,t) + \frac{\partial \sigma_i(v,t)}{\partial t} \int_v^t \sigma_i(v,y)\, dy \right) dv \right) dt \\
&\quad + \sum_{i=1}^{n} \int_0^t \frac{\partial \sigma_i(v,t)}{\partial t}\, d\tilde{z}_i(v)\, dt + \sum_{i=1}^{n} \sigma_i(t,t)\, d\tilde{z}_i(t)
\end{aligned} \tag{11.53}$$

Carverhill [**14**] gives the following necessary and sufficient condition for the above short-term interest rate specification to be Markovian:

Assuming $\sigma_i(t,T) \neq 0$, $T \in [0,\tau]$, $t \in [0,T]$, $i = 1,\ldots,n$, then there exist functions $g_i(\cdot)$ and $h_i(\cdot)$ such that:

$$\sigma_i(t,T) = g_i(t)h_i(T) \qquad \forall\, T \in [0,\tau],\ t \in [0,T],\ i = 1,\ldots,n \tag{11.54}$$

The HL model, as examined in §11.7, satisfies the above condition. It corresponds to a forward rate volatility $\sigma(t,T) = \sigma$, where σ is a strictly positive constant. The short-term interest rate is defined in (11.39) and leads to short-term interest rate dynamics of the form[23]:

$$dr(t) = \Big(\frac{\partial f(0,t)}{\partial t} + \sigma^2 t \Big)\, dt + \sigma\, d\tilde{z}(t) \tag{11.55}$$

The disadvantage, of the HL model, that only parallel shifts in the term structure are possible, is clearly demonstrated by the short-term interest rate dynamics. We expect shorter maturity forward rates to be more volatile than longer maturity forward rates. Hence, in an attempt to obtain more realistic term structure dynamics, let the forward rate volatility be a decreasing function of time, more specifically, allow for an exponentially dampened volatility

[23] The short-term interest rate dynamics may be determined using (11.53) with $n = 1$ and $\sigma(t,T) = \sigma$.

structure: $\sigma(t,T) = \sigma e^{-\gamma(T-t)}$ where γ, $\sigma > 0$ are constants. Making use of (11.32) the corresponding forward rate process is:

$$\begin{aligned} df(t,T) &= \sigma e^{-\gamma(T-t)} \int_t^T \sigma e^{-\gamma(y-t)} dy\, dt + \sigma e^{-\gamma(T-t)}\, d\tilde{z}(t) \\ &= -\frac{\sigma^2}{\gamma}\, e^{-\gamma(T-t)}\big(e^{-\gamma(T-t)} - 1\big)\, dt + \sigma e^{-\gamma(T-t)}\, d\tilde{z}(t) \end{aligned}$$

From (11.52) the short-term interest rate is expressed as:

$$\begin{aligned} r(t) &= f(0,t) + \int_0^t \frac{\sigma^2}{\gamma}\, e^{-\gamma(t-v)}\big(1 - e^{-\gamma(t-v)}\big)\, dv + \int_0^t \sigma e^{-\gamma(t-v)}\, d\tilde{z}(v) \\ &= f(0,t) + \frac{\sigma^2}{2\gamma^2}\big(1 - e^{-\gamma t}\big)^2 + \int_0^t \sigma e^{-\gamma(t-v)}\, d\tilde{z}(v) \\ &= \psi(t) + \int_0^t \sigma e^{-\gamma(t-v)}\, d\tilde{z}(v) \end{aligned}$$

where $\psi(t) = f(0,t) + \frac{\sigma^2}{2\gamma^2}\big(1 - e^{-\gamma t}\big)^2$.

Now, using (11.53), the short-term interest rate dynamics are specified as:

$$\begin{aligned} dr(t) &= \left(\frac{\partial f(0,t)}{\partial t} + \int_0^t \sigma^2 e^{-2\gamma(t-v)} - \sigma^2 e^{-\gamma(t-v)}\big(1 - e^{-\gamma(t-v)}\big) dv\right) dt \\ &\quad - \int_0^t \sigma\gamma e^{-\gamma(t-v)}\, d\tilde{z}(v)\, dt + \sigma\, d\tilde{z}(t) \\ &= \left(\frac{\partial f(0,t)}{\partial t} + \frac{\sigma^2}{\gamma}\, e^{-\gamma t}\big(1 - e^{-\gamma t}\big) - \int_0^t \sigma\gamma e^{-\gamma(t-v)}\, d\tilde{z}(v)\right) dt + \sigma\, d\tilde{z}(t) \\ &= \big(\psi'(t) + \gamma\psi(t) - \gamma r(t)\big)\, dt + \sigma\, d\tilde{z}(t) \end{aligned}$$

This is an extension of the constant parameter Vasicek model studied in Chapter 1. Here the mean reversion level $\psi'(t)/\gamma + \psi(t)$, is a function of time, but the speed of mean reversion γ, and volatility σ, are constant. The Hull–White extension of the Vasicek model (see (7.4) Chapter 7) allows for time-dependent reversion level, reversion speed and volatility. However, Hull–White observe that allowing all three parameters to be time-dependent leads to unreasonable evolution of the volatility structure and hence they recommend a model where reversion speed and volatility are kept constant. It appears that the formulation of the Vasicek model within the HJM framework results in an optimal extension of the model.

11.9.2. More general framework. Jeffrey [**32**] develops a more general result, allowing a stochastic volatility structure. However, he restricts the volatility to be a function of valuation time t, maturity time T and the time t short-term interest rate $r(t)$ only. Consequently, within a one-factor framework ($n = 1$) the HJM forward rate volatility may be expressed as $\sigma(\omega, t, T) \equiv \sigma(r, t, T)$. Notice that $\sigma(r, t, t)$ is the volatility of the short-term interest rate, as shown in (11.53).

In this analysis, the requirement that the term structure be a function of t, T and the Markovian short-term interest rate $r(t)$ leads to restrictions on the form of the forward rate volatility structure and the initial forward rate curve. These restrictions take the form of necessary and sufficient conditions indicating:

(1) which HJM models exhibit Markovian characteristics, and
(2) what volatility and initial term structure restrictions are inherent in models within a Markovian setting.

Jeffrey shows that a forward rate volatility $\sigma(r, t, T)$, leads to a Markovian short-term interest rate if there exist functions $\theta(r, t)$ and $h(t, T)$ such that:

$$\begin{aligned}(11.56)\quad \sigma(r,t,T)\int_t^T \sigma(r,t,v)\,dv = \frac{\sigma(r,t,T)}{\sigma(r,t,t)}\,\theta(r,t) + \frac{\partial}{\partial t}\left[\int_0^r \frac{\sigma(m,t,T)}{\sigma(m,t,t)}\,dm\right] \\ + h(t,T) + \tfrac{1}{2}\,\sigma(r,t,t)^2\,\frac{\partial}{\partial t}\left[\frac{\sigma(r,t,T)}{\sigma(r,t,t)}\right]\end{aligned}$$

To illustrate this condition, consider a constant forward rate volatility structure, that is $\sigma(r, t, T) = \sigma$. By (11.56) we must find a $\theta(r, t)$ and $h(t, T)$ such that:

$$\theta(r,t) = \sigma^2(T-t) - h(t,T)$$

A simple choice of $h(t, T) = \sigma^2 T$ gives $\theta(r, t) = -\sigma^2 t$ and the required condition is satisfied. Hence the associated short-term interest rate is Markovian. Now, consider a forward rate volatility structure $\sigma(r, t, T) = \sigma r^\beta$ for some $\beta > 0$. Evaluating (11.56), we require

$$\theta(r,t) = \sigma^2 r^{2\beta}(T-t) - h(t,T)$$

which is impossible, since there is no choice of $h(t, T)$ such that the right-hand side is independent of T. We may conclude that this forward rate structure is not admissible in a Markovian short-term interest rate paradigm.

Further to the above condition, Jeffrey formulates a requirement on the structure of the initial forward rate curve. Given a forward rate volatility structure satisfying (11.56), the associated initial forward rate curve must have the form:

$$(11.57)\qquad f(r,0,T) = \int_0^r \frac{\sigma(m,0,T)}{\sigma(m,0,0)}\,dm + k(T)$$

where $k(T) = -\int_0^T h(s,T)\,ds$ for any $h(t, T)$ that is valid for (11.56) to hold. Since the set of allowable initial forward rate curves is determined by the choice of $h(t,T)$ in (11.56), Jeffrey presents a result [**32**, Theorem 1] detailing the restrictions on the initial forward rate term structure.

We have $\sigma(r,t,T)$, a volatility structure satisfying (11.56) and let $\xi(t,T)$ be a deterministic function of t and T. Then one of the following is true:

- If $\sigma(r,t,T)$ is not of the form $\xi(t,T)\sigma(r,t,t)$, then there is only one pair of functions $\theta(r,t)$ and $h(t,T)$ such that (11.56) holds. Here $k(T)$ and the allowable initial forward rate curve and completely defined by $\sigma(r,t,T)$.
- If $\sigma(r,t,T)$ is of the form $\xi(t,T)\sigma(r,t,t)$, then the set of valid pairs of functions $\theta(r,t)$ and $h(t,T)$ satisfying (11.56) may be represented as:
$$\theta(r,t) = -\left[\frac{\partial}{\partial t}\int_0^r \frac{\sigma(m,t,T)}{\sigma(m,t,t)}\,dm\right]\Bigg|_{T=t} - c(t)$$
$$h(t,T) = \xi(t,T)\big(c(t) - h_p(t,t)\big) + h_p(t,T)$$
where $h_p(t,T)$ is any particular $h(t,T)$ satisfying (11.56) and $c(t)$ is any function of time t. The choice of $c(t)$ is such that any function $k(T)$ (and hence any initial forward rate curve) may be fitted in (11.57).

This result clearly demonstrates that the volatility structure and initial forward rate curve cannot be fitted independently. If the volatility structure is not of the form $\xi(t,T)\sigma(r,t,t)$, then it uniquely determines the allowable initial forward rate curve. If the volatility structure is of the form $\xi(t,T)\sigma(r,t,t)$, then $k(T)$ may be chosen to fit any initial forward rate curve.

Consider the short-term interest rate model proposed by CIR [**18**] as studied in Chapter 2. The short-term interest rate process has the form (see equation (2.17) Chapter 2):

$$dr = \kappa(\theta(t) - r)dt + \sigma\sqrt{r}dz(t) \tag{11.58}$$

Since the short-term interest rate volatility $\sigma(r,t,t) = \sigma\sqrt{r}$, is not purely deterministic, and it is reasonable that the associated forward rate volatility maintains this stochastic characteristic, it is not allowable within the analysis outlined by Carverhill. However, we show that the forward rate volatility associated with the CIR short-term interest rate process does in fact satisfy the criteria for a Markovian short-term interest rate, as set out by Jeffrey. Consider the analysis of the CIR model within the HJM framework in §11.8. The forward rate process is derived in (11.49) as:

$$df(t,T) = r(t)\left(\frac{\partial^2 \bar{B}(t,T)}{\partial T\partial t} - \kappa\frac{\partial \bar{B}(t,T)}{\partial T}\right)dt + \frac{\partial \bar{B}(t,T)}{\partial T}\,\sigma\sqrt{r(t)}\,dz(t) \tag{11.59}$$

and so the forward rate volatility $\sigma(r,t,T)$ may be found as:

$$\begin{aligned}\sigma(r,t,T) &= \frac{\partial \bar{B}(t,T)}{\partial T}\,\sigma\sqrt{r(t)} \\ &= \frac{4\gamma^2 e^{\gamma(T-t)}}{\big((\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma\big)^2}\,\sigma\sqrt{r(t)}\end{aligned}$$

Since $\sigma(r,t,T)$ is of the form $\xi(t,T)\sigma(r,t,t)$, with

$$\xi(t,T) = \frac{4\gamma^2 e^{\gamma(T-t)}}{\big((\gamma+\lambda+\kappa)(e^{\gamma(T-t)}-1)+2\gamma\big)^2}$$

we conclude that the short-term interest rate within the CIR model is Markovian and that any initial forward rate curve may be fitted[24].

11.10. Conclusion

HJM develop a new methodology for modelling the term structure of interest rates. They make use of a process describing the evolution forward rates to derive a methodology for contingent claim valuation, which is free from arbitrage and independent of the market prices of risk. By modelling forward rates, the stochastic behaviour of the entire term structure, not just the short-term interest rate, is modelled at any point in time. This allows information from the term structure to be used to eliminate the dependence on market prices of risk.

For the single factor case, the HJM formulation does not add much to previously developed models such as the Hull–White extended Vasicek and BDT models. In fact, the complexity of the calibration and contingent claim valuation procedures may act as a deterrent. However, within a multi-factor context the elegance of the HJM framework is undeniable. The methodology provides a coherent framework allowing easy incorporation of additional factors. The resulting increase in computational time tends to be linear (as opposed to exponential increases exhibited by other models). This is because the non-Markovian nature of the model makes Monte Carlo simulation the valuation technique of choice. This allows easy valuation of path-dependent options, but does become problematic for American-style contingent claims.

[24]The admissible initial forward rate curve will be subjected to the requirement that $2\kappa\theta(t) \geq \sigma^2$, as explained in §11.8. However, this restriction is required to prevent negative interest rates and is not associated with the Markovian nature of the short-term interest rate process.

CHAPTER 12

Brace, Gatarek and Musiela Model

12.1. Introduction

12.1.1. A continuum of forward rates. All the models examined thus far have been based on instantaneous short-term or forward interest rates. This implies that the fundamental building blocks, that is default-free bonds, are assumed to be continuous (or smooth) with respect to the tenor. Even the discrete time models such as Ho and Lee [**27**] (see Chapter 10) and Black, Derman and Toy [**6**] (see Chapter 8), which make use of a discrete set of discount bonds, assume these are extracted from an underlying continuum of default-free bonds. Such a continuum of default-free discount bonds is not actually traded, nor does the associated continuum of instantaneous short-term or forward interest rates exist.

This assumption need not be problematic, since calibration of the models often requires a discretisation of the continuous time processes. Additionally, traded instruments are only contingent on a discrete number of points on the yield curve. For example, the pricing and hedging of a forward contract on the discrete forward rate[1] $f(t, T, T+\delta)$ requires the existence of two bonds $P(t, T)$ and $P(t, T+\delta)$, maturing at the expiry and payoff times respectively. Similarly, a swap-based product depends on bonds maturing at the start of the swap and at the payment times of the fixed leg. Usually it is only a small set of discrete discount bonds that determines the price and associated hedge of such LIBOR-based[2] instruments. Given a complete set of spanning forward rates, the required set of bonds may be recovered. Hence a complete set of spanning forward rates provides a sufficient description of the yield curve enabling the pricing of LIBOR-based instruments.

12.1.2. The lognormality assumption. Caps and floors are fundamental components within a swap and swap derivative market. A cap (floor) is a strip of caplets (floorlets) which are calls (puts) on an underlying forward rate. The market convention is to assume a lognormal structure for the forward rate process and hence to price each of these options using the Black

[1]Here $f(t, T, T+\delta)$ represents the time t value of the forward rate applicable over the interval $[T, T+\delta]$.

[2]London Interbank Offer Rate. This is one of the most frequently used discrete forward rates.

futures formula. However, as shown in §11.5, allowing the instantaneous forward rate to assume a lognormal volatility structure causes it to explode in finite time. This implies that all forward rates cannot be lognormal under a single arbitrage-free measure. One could conclude that the market prices of caps and floors are flawed in some way and inconsistent with an arbitrage-free framework.

Brace, Gaterek and Musiela (BGM) [**9**] consider discretely compounded forward rates and show that a lognormal structure may be imposed while maintaining an arbitrage-free framework. In the HJM model of instantaneous forward rates, a single spot arbitrage-free measure is applied to all forward rates; while BGM assign, to each forward rate, a forward arbitrage-free measure defined by the settlement date of the associated forward rate. This model then justifies the use of the Black futures formula for pricing caps and floors[3].

12.2. Initial framework

The mathematical framework used by BGM is much the same as that used by HJM. We have a probability space $(\Omega, \{F_t; t \geq 0\}, \widetilde{Q})$ where Ω is the state space and the filtration $\{F_t; t \geq 0\}$ is the $\widetilde{Q}$-augmentation of the filtration generated by the n-dimensional Brownian motion $\tilde{z} = \{\tilde{z}(t); t \geq 0\}$. Since we make use of the arbitrage-free results of the HJM analysis, $\widetilde{Q}$ is the risk-neutral probability measure with $\tilde{z}$ the corresponding Brownian motion. The trading interval is specified as $[0, \tau]$ where $\tau > 0$ is fixed. The following processes are defined on this probability space[4]:

- $f(t,T)$ denotes the instantaneous, continuously compounded forward rate prevailing at time t for maturity T. The process $\{f(t,T); t \leq T\}$ satisfies:

$$df(t,T) = \sigma(t,T) \cdot \sigma^*(t,T)\, dt + \sigma(t,T) \cdot d\tilde{z}(t) \tag{12.1}$$

  where $\sigma(t,T)$ is the forward rate volatility and $\sigma^*(t,T) = \int_t^T \sigma(t,v)\, dv$.
- $P(t,T) = \exp\left(-\int_t^T f(t,u)du\right)$ describes the price evolution of a T-maturity discount bond, and so:

$$dP(t,T) = P(t,T)r(t)\, dt - P(t,T)\sigma^*(t,T) \cdot d\tilde{z}(t) \tag{12.2}$$

[3]Here the observation needs to be made that the market does not appear to distinguish between forward measures, and hence forward probabilities, at different maturities

[4]The initial BGM formulation [**9**] of these processes uses $r(t,x)$ to represent the instantaneous forward rate prevailing at time t for maturity $t+x$. I feel this formulation obscures any value it adds and hence I maintain consistency with the notation used in Chapter 11 by using $f(t,T)$ to denote the time t instantaneous forward rate for maturity T. The obvious relationship between the two representations is $r(t,x) = f(t,t+x)$.

where $\sigma^*(t,T)$ may be interpreted as bond price volatility and hence $\sigma^*(t,t) = 0$ for all $t \geq 0$.

- Defining the short-term interest rate $r(t) = f(t,t)$ for all $t \geq 0$, the money market account is represented as:

$$B(t) = \exp\left(\int_0^t r(v)\,dv\right) \tag{12.3}$$

with initial condition $B(0) = 1$.

We know that if discounted bond prices $\frac{P(t,T)}{B(t)}$, $t \in [0,T]$, $T > 0$ are martingales under some probability measure $\widetilde{Q}$, then we are in an arbitrage-free framework. Within this framework, the bond price may be represented as[5]:

$$\frac{P(t,T)}{B(t)} = \frac{P(0,T)}{B(0)} \exp\left(-\int_0^t \sigma^*(s,T)\cdot d\tilde{z}(s) - \frac{1}{2}\int_0^t \left|\sigma^*(s,T)\right|^2\,ds\right) \tag{12.4}$$

12.3. Model of the forward LIBOR rate

Specifying the above model of the instantaneous, continuously compounded forward rate is equivalent to determining the volatility function $\sigma(t,T)$ (or equivalently $\sigma^*(t,T)$). Fix some $\delta > 0$, then the LIBOR rate process $\{L(t,T);\ t \in [0,T], T \in [0,\tau]\}$ is defined as:

$$1 + \delta L(t,T) = \exp\left(\int_T^{T+\delta} f(t,v)\,dv\right) \tag{12.5}$$

Imposing a lognormal volatility structure on $L(t,T)$, the associated stochastic process may be written as:

$$dL(t,T) = \mu_{L(t,\,T)}\,dt + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t) \tag{12.6}$$

where $\mu_{L(t,\,T)}$ is some drift function and $\gamma : R^2 \to R^n$ is the deterministic, bounded and piecewise continuous relative volatility function. Letting $h(t,T) = \int_T^{T+\delta} f(t,u)\,du$, we make use of Ito's Lemma to determine the correct functional form of (12.6). Hence:

$$dL(t,T) = \frac{\partial L(t,T)}{\partial h(t,T)}\,dh(t,T) + \frac{1}{2}\frac{\partial^2 L(t,T)}{\partial h(t,T)^2}\,dh(t,T)\,dh(t,T) \tag{12.7}$$

[5]See equation (11.28).

Here:

$$
\begin{aligned}
dh(t,T) &= d\left(\int_T^{T+\delta} f(t,u)\,du\right) \\
&= \int_T^{T+\delta} df(t,u)\,du \\
&= \int_T^{T+\delta}\Big(\sigma(t,u)\cdot\sigma^*(t,u)\,dt + \sigma(t,u)\cdot d\tilde{z}(t)\Big)du \\
&= \int_T^{T+\delta} \frac{1}{2}\frac{\partial\sigma^*(t,u)^2}{\partial u}\,du\,dt + \int_T^{T+\delta}\sigma(t,u)\,du\cdot d\tilde{z}(t) \\
&= \frac{1}{2}\Big(\left|\sigma^*(t,T+\delta)\right|^2 - \left|\sigma^*(t,T)\right|^2\Big)\,dt + \int_t^{T+\delta}\sigma(t,u)\,du\cdot d\tilde{z}(t) \\
&\quad - \int_t^{T}\sigma(t,u)\,du\cdot d\tilde{z}(t) \\
&= \frac{1}{2}\Big(\left|\sigma^*(t,T+\delta)\right|^2 - \left|\sigma^*(t,T)\right|^2\Big)\,dt \\
&\quad + \Big(\sigma^*(t,T+\delta) - \sigma^*(t,T)\Big)\cdot d\tilde{z}(t)
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
dh(t,T)\,dh(t,T) &= d\left(\int_T^{T+\delta} f(t,u)\,du\right) d\left(\int_T^{T+\delta} f(t,u)\,du\right) \\
&= \left|\sigma^*(t,T+\delta) - \sigma^*(t,T)\right|^2\,dt
\end{aligned}
$$

and (12.7) becomes:

$$
\begin{aligned}
dL(t,T) &= \frac{1}{\delta}\exp\left(\int_T^{T+\delta} f(t,u)\,du\right)\left(\frac{1}{2}\Big(\left|\sigma^*(t,T+\delta)\right|^2 - \left|\sigma^*(t,T)\right|^2\Big)\,dt\right. \\
&\quad + \frac{1}{2}\left|\sigma^*(t,T+\delta) - \sigma^*(t,T)\right|^2\,dt \\
&\quad \left. + \Big(\sigma^*(t,T+\delta) - \sigma^*(t,T)\Big)\cdot d\tilde{z}(t)\right) \\
&= \frac{1}{\delta}\Big(1+\delta L(t,T)\Big)\Big(\left|\sigma^*(t,T+\delta)\right|^2 - \sigma^*(t,T)\cdot\sigma^*(t,T+\delta)\Big)dt \\
&\quad + \frac{1}{\delta}\Big(1+\delta L(t,T)\Big)\Big(\sigma^*(t,T+\delta) - \sigma^*(t,T)\Big)\cdot d\tilde{z}(t) \\
&= \frac{1}{\delta}\Big(1+\delta L(t,T)\Big)\sigma^*(t,T+\delta)\cdot\Big(\sigma^*(t,T+\delta) - \sigma^*(t,T)\Big)dt \qquad (12.8) \\
&\quad + \frac{1}{\delta}\Big(1+\delta L(t,T)\Big)\Big(\sigma^*(t,T+\delta) - \sigma^*(t,T)\Big)\cdot d\tilde{z}(t)
\end{aligned}
$$

Hence by (12.6) we require:

$$\begin{aligned}
&\frac{1}{\delta}\Big(1+\delta L(t,T)\Big)\Big(\sigma^*(t,T+\delta)-\sigma^*(t,T)\Big) = L(t,T)\gamma(t,T) \\
&\Rightarrow \sigma^*(t,T+\delta)-\sigma^*(t,T) = \frac{\delta L(t,T)}{1+\delta L(t,T)}\,\gamma(t,T)
\end{aligned}\tag{12.9}$$

and so (12.8) may be written in terms of the $(T+\delta)$-maturity bond price volatility as:

$$dL(t,T) = L(t,T)\gamma(t,T)\cdot\sigma^*(t,T+\delta)dt + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t)\tag{12.10}$$

Alternatively, solving (12.9) for $\sigma^*(t,T+\delta)$ we may write this LIBOR stochastic process in terms of the T-maturity bond price volatility as:

$$\begin{aligned}
dL(t,T) = \Big(&L(t,T)\gamma(t,T)\cdot\sigma^*(t,T) \\
&+\frac{\delta L(t,T)^2}{1+\delta L(t,T)}\,|\gamma(t,T)|^2\Big)dt + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t)
\end{aligned}\tag{12.11}$$

Now let us assume[6] $\sigma^*(t,T)=0$ for all $t\in\Big((T-\delta)\vee 0, T\Big]$ and $T\in[0,\tau]$, then a recursive relationship may be used to define $\sigma^*(t,T)$ for $T-t\geq\delta$ as[7]:

$$\sigma^*(t,T) = \sum_{k=1}^{\delta^{-1}(T-t)} \frac{\delta L(t,T-k\delta)}{1+\delta L(t,T-k\delta)}\,\gamma(t,T-k\delta)\tag{12.12}$$

[6]This assumption implies the volatility factor disappears for all rates where $0\leq T-t<\delta$, that is the time between valuation date and maturity date is less than δ. This allows for the construction of a tractable model. We have $\sigma^*(t,t)=0$ for all $t\in[0,T]$ since this is the price volatility of an instantly maturing bond. Relationship (12.9) implies:

$$\sigma^*(t,T)=\sigma^*(t,T-\delta)+\frac{\delta L(t,T-\delta)}{1+\delta L(t,T-\delta)}\,\gamma(t,T-\delta)$$

Hence, for $T=t+\delta$

$$\sigma^*(t,t+\delta)=\sigma^*(t,t)+\frac{\delta L(t,t)}{1+\delta L(t,t)}\,\gamma(t,t)=0$$

since $\gamma(t,t)=0$ is the volatility of the spot LIBOR rate. So, since $\sigma^*(t,T)=0$ for $T=t$ and for $T=t+\delta$ we let $\sigma^*(t,T)=0$ for all $T\in(t,t+\delta)$ as well. This is equivalent to $\sigma^*(t,T)=0$ for $t\in(T-\delta,T)$.

Substituting this recursive relationship into (12.10), the stochastic process describing the evolution of LIBOR may be written purely in terms of LIBOR rate volatilities as:

$$\begin{aligned} dL(t,T) = L(t,T)\gamma(t,T)\cdot \sum_{k=1}^{\delta^{-1}(T-t)} \frac{\delta L(t,T+\delta-k\delta)}{1+\delta L(t,T+\delta-k\delta)}\gamma(t,T+\delta-k\delta)\,dt \\ + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t) \end{aligned}$$

Letting $j = k-1$ we have:

$$\begin{aligned} dL(t,T) = L(t,T)\gamma(t,T)\cdot \sum_{j=0}^{\delta^{-1}(T-t)-1} \frac{\delta L(t,T-j\delta)}{1+\delta L(t,T-j\delta)}\gamma(t,T-j\delta)\,dt \\ + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t) \end{aligned} \tag{12.13}$$

12.4. Forward risk-neutral measure

The framework within which the HJM model is derived, in fact the framework within which the above analysis is performed, is the risk-neutral framework. This is characterised by a risk-neutral probability measure under which all discounted asset prices are martingales. Other risk-neutral probability measures may be defined in a similar way, allowing various asset/option pricing problems to be solved. Changing between probability measures is associated with numeraire changes. This allows the numeraire to be chosen in such as way as to expedite the option valuation procedure.

Consider a non-dividend-paying security $X(t)$, which is a martingale under the risk-neutral probability measure $\widetilde{Q}$. There exists a probability measure Q^X such that the price of any asset relative to X is a Q^X-martingale. The probability measure Q^X is defined by the Radon–Nikodym derivative with respect to $\widetilde{Q}$ as [**22**, Theorem 1]:

[7] By (12.9) we have:

$$\begin{aligned} \sigma^*(t,T) &= \sigma^*(t,T-\delta) + \frac{\delta L(t,T-\delta)}{1+\delta L(t,T-\delta)}\gamma(t,T-\delta) \\ &= \sigma^*(t,T-2\delta) + \frac{\delta L(t,T-2\delta)}{1+\delta L(t,T-2\delta)}\gamma(t,T-2\delta) + \frac{\delta L(t,T-\delta)}{1+\delta L(t,T-\delta)}\gamma(t,T-\delta) \\ &= \sigma^*(t,T-k\delta) + \frac{\delta L(t,T-k\delta)}{1+\delta L(t,T-k\delta)}\gamma(t,T-k\delta) + \cdots \\ &\quad + \frac{\delta L(t,T-\delta)}{1+\delta L(t,T-\delta)}\gamma(t,T-\delta) \end{aligned}$$

Since, by assumption, $\sigma^*(t,T) = 0$ for $T-t < \delta$ the first term on the RHS vanishes for $T-t-k\delta < \delta$ i.e. the term vanishes for $k > \delta^{-1}(T-t) - 1$, hence the upper bound for the summation index is $k = \delta^{-1}(T-t)$.

$$\frac{dQ^X}{d\widetilde{Q}} = \frac{X(T)}{X(0)B(T)} \tag{12.14}$$

since the money market account $B(t)$, is the numeraire asset associated with probability measure $\widetilde{Q}$.

Calculating the time t price of an asset generating a cash flow at the future time T lends itself to a T-maturity discount bond as numeraire. Let Q^T denote the probability measure associated with $P(t,T)$ (T-maturity discount bond) as numeraire. Making use of (12.14), we define this probability measure as:

$$\begin{aligned} \frac{dQ^T}{d\widetilde{Q}} &= \frac{P(T,T)}{P(0,T)B(T)} \\ &= \exp\left(-\int_0^T \sigma^*(s,T)\cdot d\tilde{z}(s) - \frac{1}{2}\int_0^T |\sigma^*(s,T)|^2\, ds\right) \end{aligned} \tag{12.15}$$

Now since the relative price of a T-maturity discount bond is a martingale with respect to the risk-neutral measure and $P(0,T)$ is known at time t, we define:

$$\begin{aligned} \eta_t &= \left.\frac{dQ^T}{d\widetilde{Q}}\right|_{F_t} \\ &= \mathbb{E}^{\widetilde{Q}}\left[\frac{P(T,T)}{P(0,T)B(T)}\,\middle|\, F_t\right] \\ &= \frac{P(t,T)}{P(0,T)B(t)} \\ &= \exp\left(-\int_0^t \sigma^*(s,T)\cdot d\tilde{z}(s) - \frac{1}{2}\int_0^t |\sigma^*(s,T)|^2\, ds\right) \end{aligned}$$

By the Radon–Nikodym derivative (12.15), we have for any asset $S(\cdot)$:

$$\mathbb{E}^{Q^T}[S(T)] = \mathbb{E}^{\widetilde{Q}}\left[S(T)\,\frac{dQ^T}{d\widetilde{Q}}\right] = \mathbb{E}^{\widetilde{Q}}[S(T)\,\eta_T] \tag{12.16}$$

and for conditional expectations we make use of Bayes' Rule to yield [**49**]:

$$\begin{aligned} \mathbb{E}^{Q^T}[S(T)\,|\,F_t] &= \frac{1}{\eta_t}\,\mathbb{E}^{\widetilde{Q}}[S(T)\,\eta_T\,|\,F_t] \\ &= \frac{B(t)P(0,T)}{P(t,T)}\,\mathbb{E}^{\widetilde{Q}}\left[\frac{S(T)}{B(T)P(0,T)}\,\middle|\, F_t\right] \\ &= \frac{S(t)}{P(t,T)} \end{aligned} \tag{12.17}$$

since the relative asset price $\frac{S(\cdot)}{B(\cdot)}$ is a martingale under probability measure $\widetilde{Q}$ and $P(0,T)$ is known at time t. Since the time t forward price for time T of asset S may be written as [**41**]:

$$F_{S_T}(t,T) = \frac{S(t)}{P(t,T)}$$

we observe that the forward price process $F_{S_T}(t,T)$, $t \in [0,T]$, follows a martingale[8] under the measure Q^T. Hence this measure is referred to as the forward martingale measure or forward neutral probability measure.

We use Girsanov's Theorem (for example see [**41**, Theorem B.2.1.]) to define the new Brownian motion under the forward martingale measure as:

$$z^T(t) = \tilde{z}(t) + \int_0^t \sigma^*(u,T)\,du \tag{12.18}$$

12.5. Forward LIBOR rate with respect to the forward measure

Consider (12.10), the LIBOR rate process represented in terms of the $(T+\delta)$-maturity bond price volatility:

$$dL(t,T) = L(t,T)\gamma(t,T)\cdot\sigma^*(t,T+\delta)dt + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t) \tag{12.19}$$

We may introduce a new n-dimensional process $z^{T+\delta}(t)$ corresponding to time $T+\delta$ where

$$dz^{T+\delta}(t) = d\tilde{z}(t) + \sigma^*(t,T+\delta)dt \tag{12.20}$$

and a corresponding probability measure $Q^{T+\delta}$, equivalent to $\widetilde{Q}$, under which $z^{T+\delta}(t)$ is a Brownian motion. By (12.15), this new probability measure, called the forward measure, may be defined as:

$$\begin{aligned}\frac{dQ^{T+\delta}}{d\widetilde{Q}} &= \frac{P(T+\delta,T+\delta)}{P(0,T+\delta)B(T+\delta)}\\ &= \exp\left(-\int_0^{T+\delta}\sigma^*(s,T+\delta)\cdot d\tilde{z}(s) - \frac{1}{2}\int_0^{T+\delta}\left|\sigma^*(s,T+\delta)\right|^2 ds\right) \qquad (12.21)\end{aligned}$$

[8]Since

$$F_{S_T}(T,T) = \frac{S(T)}{P(T,T)} = S(T)$$

we may write

$$\mathbb{E}^{Q^T}\left[F_{S_T}(T,T)\,|\,F_t\right] = \mathbb{E}^{Q^T}\left[S(T)\,|\,F_t\right] = \frac{S(t)}{P(t,T)} = F_{S_T}(t,T)$$

and hence $F_{S_T}(t,T)$, $t\in[0,T]$ follows a martingale.

Now, considering (12.19) and making use of (12.20) we have:

$$\begin{aligned} dL(t,T) &= L(t,T)\gamma(t,T)\cdot\sigma^*(t,T+\delta)dt + L(t,T)\gamma(t,T)\cdot d\tilde{z}(t) \\ &= L(t,T)\gamma(t,T)\cdot\Big(d\tilde{z}(t)+\sigma^*(t,T+\delta)dt\Big) \\ &= L(t,T)\gamma(t,T)\cdot dz^{T+\delta}(t) \end{aligned} \tag{12.22}$$

Hence each forward LIBOR rate $L(t,T)$ follows a lognormal martingale process under the forward measure corresponding to its settlement date $T+\delta$.

Using a forward measure for a date other than the settlement date, will require a drift adjustment. Consider (12.11), the LIBOR rate process expressed in terms of the T-maturity bond price volatility:

$$\begin{aligned} dL(t,T) = &\Big(L(t,T)\gamma(t,T)\cdot\sigma^*(t,T) + \frac{\delta L(t,T)^2}{1+\delta L(t,T)}\,|\gamma(t,T)|^2\Big)dt \\ &+ L(t,T)\gamma(t,T)\cdot d\tilde{z}(t) \end{aligned} \tag{12.23}$$

Making use of (12.20) and (12.21), we define the Brownian motion and forward probability measure corresponding to time T, the expiry date of the forward LIBOR rate, as:

$$dz^T(t) = d\tilde{z}(t) + \sigma^*(t,T)dt \tag{12.24}$$

and

$$\begin{aligned} \frac{dQ^T}{d\widetilde{Q}} &= \frac{P(T,T)}{P(0,T)B(T)} \\ &= \exp\left(-\int_0^T \sigma^*(s,T)\cdot d\tilde{z}(s) - \frac{1}{2}\int_0^T |\sigma^*(s,T)|^2\, ds\right) \end{aligned} \tag{12.25}$$

respectively, and so (12.23) becomes:

$$\begin{aligned} dL(t,T) &= \frac{\delta L(t,T)^2}{1+\delta L(t,T)}\,|\gamma(t,T)|^2\,dt + L(t,T)\gamma(t,T)\cdot\Big(d\tilde{z}(t)+\sigma^*(t,T)\,dt\Big) \\ &= \frac{\delta L(t,T)^2}{1+\delta L(t,T)}\,|\gamma(t,T)|^2\,dt + L(t,T)\gamma(t,T)\cdot dz^T(t) \end{aligned}$$

where $\frac{\delta L(t,T)^2}{1+\delta L(t,T)}\,|\gamma(t,T)|^2$ is the drift adjustment required when a forward measure corresponding to the expiry date of LIBOR is used. Making use of the recursive relationship of bond price volatilities shown in (12.9) and (12.12), the drift adjustment for any forward measure may be found. Equation (12.13) shows the drift adjustment when the spot measure is used. This corresponds to the money market account as numeraire and hence may be viewed as the time t measure. From this we may conclude that under this spot measure, no forward LIBOR rate follows a lognormal martingale.

12.6. Derivative pricing

Let us consider a cap on the forward LIBOR rate. Settlement is in arrears at times T_j, $j = 1, \ldots, n$, with cash flows at time T_j equal to $\delta(L(T_{j-1}, T_{j-1}) -\kappa)^+$ for each $j = 1, \ldots, n$, where κ is the strike price of the cap. Consider the cap price at time t, $t \leq T_0$. By the risk-neutral valuation principle, the arbitrage-free price is given by:

$$\frac{\mathrm{Cap}(t)}{B(t)} = \sum_{j=1}^{n} \mathbb{E}^{\widetilde{Q}} \left[\left. \frac{\delta\Big(L(T_{j-1}, T_{j-1}) - \kappa\Big)^+}{B(T_j)} \right| F_t \right] \tag{12.26}$$

From (12.17) we have:

$$\mathbb{E}^{\widetilde{Q}} \left[S(T) \,|\, F_t \right] = \eta_t \, \mathbb{E}^{Q^T} \left[\left. \frac{S(T)}{\eta_T} \right| F_t \right]$$

and so:

$$\begin{aligned}
&\mathbb{E}^{\widetilde{Q}} \left[\left. \frac{\delta\Big(L(T_{j-1}, T_{j-1}) - \kappa\Big)^+}{B(T_j)} \right| F_t \right] \\
&= \frac{P(t, T_j)}{P(0, T_j) B(t)} \, \mathbb{E}^{Q^{T_j}} \left[\left. \frac{\delta\Big(L(T_{j-1}, T_{j-1}) - \kappa\Big)^+}{B(T_j)} P(0, T_j) B(T_j) \right| F_t \right] \\
&= \frac{P(t, T_j)}{B(t)} \, \mathbb{E}^{Q^{T_j}} \left[\left. \delta\Big(L(T_{j-1}, T_{j-1}) - \kappa\Big)^+ \right| F_t \right]
\end{aligned}$$

Substituting into (12.26) the cap price may be written as:

$$\mathrm{Cap}(t) = \sum_{j=1}^{n} P(t, T_j) \, \mathbb{E}^{Q^{T_j}} \left[\left. \delta\Big(L(T_{j-1}, T_{j-1}) - \kappa\Big)^+ \right| F_t \right] \tag{12.27}$$

where each cash flow is valued under the forward measure corresponding to payment date T_j. Here $\mathbb{E}^{Q^{T_j}}[\cdot]$ denotes expectation under the time T_j forward measure Q^{T_j}. This forward measure is defined, using (12.15), as:

$$
\begin{aligned}
\frac{dQ^{T_j}}{d\widetilde{Q}} &= \Big(P(0,T_j)B(T_j)\Big)^{-1} \\
&= \exp\left(-\int_0^{T_j} \sigma^*(s,T_j)\cdot d\tilde{z}(s) - \frac{1}{2}\int_0^{T_j} \left|\sigma^*(s,T_j)\right|^2\, ds\right)
\end{aligned}
$$

From (12.22) we know $L(t,T_j)$ may be represented as a lognormal martingale under the forward measure $Q^{T_{j+1}}$; hence:

$$dL(t,T_j) = L(t,T_j)\gamma(t,T_j)\cdot dz^{T_{j+1}}(t)$$

and so given the time t forward value of the time T_j LIBOR rate we have:

$$
\begin{aligned}
L(T_j,T_j) = L(t,T_j)\exp\Bigg(&-\int_t^{T_j}\gamma(s,T_j)\cdot dz^{T_{j+1}}(s) \\
&-\frac{1}{2}\int_t^{T_j}\left|\gamma(s,T_j)\right|^2\, ds\Bigg)
\end{aligned}
\tag{12.28}
$$

Now consider the process followed by $\frac{1}{L(t,T_j)}$. By Ito's Lemma we have:

$$
\begin{aligned}
d\frac{1}{L(t,T_j)} &= \frac{\partial \frac{1}{L(t,T_j)}}{\partial L(t,T_j)}\, dL(t,T_j) + \frac{1}{2}\frac{\partial^2 \frac{1}{L(t,T_j)}}{\partial L(t,T_j)^2}\, dL(t,T_j)dL(t,T_j) \\
&= -\frac{\gamma(t,T_j)}{L(t,T_j)}\, dz^{T_{j+1}}(t) + \frac{\gamma(t,T_j)^2}{L(t,T_j)}\, dt \\
&= \gamma(t,T_j)\frac{1}{L(t,T_j)}\cdot dz^L(t)
\end{aligned}
$$

where

$$dz^L(t) = -dz^{T_{j+1}}(t) + \gamma(t,T_j)\, dt$$

is the Brownian motion corresponding to probability measure Q^L under which $\frac{1}{L(t,T_j)}$ is a lognormal martingale. By the Girsanov Theorem, the Radon–Nikodym derivative defining this change of measure may be expressed as:

$$
\begin{aligned}
\frac{dQ^L}{dQ^{T_{j+1}}} &= \exp\left(-\int_0^{T_j}\gamma(s,T_j)\cdot dz^{T_{j+1}}(s) - \frac{1}{2}\int_0^{T_j}\left|\gamma(s,T_j)\right|^2\, ds\right) \\
&= \frac{L(T_j,T_j)}{L(t,T_j)}
\end{aligned}
$$

This change of measure affects the drift but not the volatility coefficient, hence as in (12.28) we may write:

$$\frac{1}{L(T_j,T_j)} = \frac{1}{L(t,T_j)} \exp\left(-\int_t^{T_j} \gamma(s,T_j)\cdot dz^L(s) - \frac{1}{2}\int_t^{T_j} |\gamma(s,T_j)|^2\, ds \right) \tag{12.29}$$

From (12.27) we see that the cap is made up of a series of n call options on LIBOR, or caplets, maturing at times T_j, $j = 1, \ldots, n$ respectively. Consider the value of a caplet with cash flow occurring at time T_{j+1}:

$$\begin{aligned}
\text{Caplet}(t) &= \delta P(t,T_{j+1})\, \mathbb{E}^{Q^{T_{j+1}}} \left[\Big(L(T_j,T_j) - \kappa \Big)^+ \,\Big|\, F_t \right] \\
&= \delta P(t,T_{j+1})\, \mathbb{E}^{Q^{T_{j+1}}} \left[\Big(L(T_j,T_j) - \kappa \Big) 1_{\{L(T_j,T_j)>\kappa\}} \,\Big|\, F_t \right] \\
&= \delta P(t,T_{j+1}) \left(L(t,T_j)\, \mathbb{E}^{Q^{T_{j+1}}} \left[\frac{L(T_j,T_j)}{L(t,T_j)}\, 1_{\{L(T_j,T_j)>\kappa\}} \,\Big|\, F_t \right] \right. \\
&\quad \left. -\kappa\, \mathbb{E}^{Q^{T_{j+1}}} \left[1_{\{L(T_j,T_j)>\kappa\}} \,\big|\, F_t \right] \right) \\
&= \delta P(t,T_{j+1}) \Big(L(t,T_j)\, \mathbb{P}^{Q^L} \Big\{ L(T_j,T_j) > \kappa \Big\} \\
&\quad -\kappa\, \mathbb{P}^{Q^{T_{j+1}}} \Big\{ L(T_j,T_j) > \kappa \Big\} \Big) \\
&= \delta P(t,T_{j+1}) \left(L(t,T_j)\, \mathbb{P}^{Q^L} \left\{ \frac{1}{L(T_j,T_j)} < \frac{1}{\kappa} \right\} \right. \\
&\quad \left. -\kappa\, \mathbb{P}^{Q^{T_{j+1}}} \Big\{ L(T_j,T_j) > \kappa \Big\} \right) \qquad (12.30)
\end{aligned}$$

Hence, making use of (12.29) we have:

$$\begin{aligned}
&\mathbb{P}^{Q^L} \left\{ \frac{1}{L(T_j,T_j)} < \frac{1}{\kappa} \right\} \\
&= \mathbb{P}^{Q^L} \left\{ \exp\left(-\int_t^{T_j} \gamma(s,T_j)\cdot dz^L(s) - \frac{1}{2}\int_t^{T_j} |\gamma(s,T_j)|^2\, ds \right) < \frac{L(t,T_j)}{\kappa} \right\} \\
&= \mathbb{P}^{Q^L} \left\{ -\int_t^{T_j} \gamma(s,T_j)\cdot dz^L(s) < \ln\left(\frac{L(t,T_j)}{\kappa} \right) + \frac{1}{2}\int_t^{T_j} |\gamma(s,T_j)|^2\, ds \right\} \\
&= N\big(h(t,T_j)\big)
\end{aligned}$$

where

$$h(t,T_j) = \frac{\ln\left(\frac{L(t,T_j)}{\kappa}\right) + \frac{1}{2}\varphi^2(t,T_j)}{\varphi(t,T_j)}$$

$$\varphi^2(t,T_j) = \int_t^{T_j} |\gamma(s,T_j)|^2 \, ds \tag{12.31}$$

Similarly, making use of (12.28), we have:

$$\begin{aligned}
&\mathbb{P}^{Q^{T_{j+1}}}\Big\{L(T_j,T_j) > \kappa\Big\} \\
&= \mathbb{P}^{Q^{T_{j+1}}}\left\{-\int_t^{T_j} \gamma(s,T_j)\cdot dz^{T_{j+1}}(s) > \ln\left(\frac{\kappa}{L(t,T_j)}\right) + \frac{1}{2}\int_t^{T_j} |\gamma(s,T_j)|^2 \, ds\right\} \\
&= \mathbb{P}^{Q^{T_{j+1}}}\left\{\int_t^{T_j} \gamma(s,T_j)\cdot dz^{T_{j+1}}(s) < \ln\left(\frac{L(t,T_j)}{\kappa}\right) - \frac{1}{2}\int_t^{T_j} |\gamma(s,T_j)|^2 \, ds\right\} \\
&= N\big(h(t,T_j) - \varphi(t,T_j)\big)
\end{aligned}$$

Substituting these results into (12.30), the time t value of a caplet with cash flow at time T_{j+1} may be written as:

$$\begin{aligned}
\text{Caplet}(t) = \delta P(t,T_{j+1})\Big(&L(t,T_j)\, N\big(h(t,T_j)\big) \\
&-\kappa\, N\big(h(t,T_j) - \varphi(t,T_j)\big)\Big)
\end{aligned} \tag{12.32}$$

and so, by equation (12.27) the time t value of the entire cap is:

$$\begin{aligned}
\text{Cap}(t) = \sum_{j=1}^{n} \delta P(t,T_j)\Big(&L(t,T_{j-1})\, N\big(h(t,T_{j-1})\big) \\
&-\kappa\, N\big(h(t,T_{j-1}) - \varphi(t,T_{j-1})\big)\Big)
\end{aligned} \tag{12.33}$$

12.7. Calibration to market volatilities

One of the main advantages of the BGM model is that calibration no longer involves the translation of unobservable state variables (such as instantaneous spot and forward rates) into quantities observed in the market. The volatility parameter $\varphi^2(t,T_j)$ in (12.31) may be directly calibrated to the observed Black volatilities. However $\varphi^2(t,T_j)$ is the volatility for a single caplet, while quoted Black cap volatilities are in fact 'average' volatilities over a series of caplets. In order to derive individual caplet volatilities from quoted Black cap volatilities, assumptions may need to be made about the relationships between individual caplet volatilities. These caplet volatilities, which represent volatilities of individual forward rates, are referred to as forward-forward volatilities.

In the derivation of the above pricing formulae, the volatility function γ is n-dimensional, allowing for n volatility factors (sources of uncertainty). These factors can be identified from market data along similar lines detailed in Chapter 15 which details calibration of the HJM approach.

The inconsistency within market caplet and swaption prices remains. All market prices of caplets and swaptions are derived from a model which simultaneously assumes a lognormal distribution for the underlying remains. While the lognormality assumption is valid for any individual caplet or swaption, joint lognormality of all caplets and swaptions is impossible. The pricing impact of this internal inconsistency is likely to be small [**45**]. While one cannot exactly calibrate a single model to all caplet and swaption volatilities several optimised hybrid approaches have been suggested (e.g. [**46**]).

12.8. Conclusion

The implications of the modelling approach presented by BGM are twofold. Firstly, we see that, in order to justify the market convention of pricing caplets by means of the Black formula, one needs to consider forward rates under a forward measure [**9**]. No forward rates are lognormal under the spot measure, but rather under an appropriate forward measure, which depends on the settlement date of the forward rate.

Secondly, the modelling framework presented here is somewhat different to that used by previously studied models. Other models, specifically the HJM model (see Chapter 11), model unobservable market parameters. Instantaneous forward rates, as modelled by HJM, are not observable and so implementation requires a suitable discretisation. BGM have developed a continuous time model of discrete forward rates which are market observable quantities.

One of the most difficult tasks faced by users of the traditional models is that of ensuring the recovery of market-observed values and volatilities [**45**]. Within the BGM framework the modelled variables are in fact the market-observed quantities, and hence one is spared the difficult task of transforming unobservable model parameters into values of traded quantities.

Another more subtle advantage is that the BGM model may be used to directly express views about future values and volatilities of market observables. Via the BGM model, these predictions are directly translated into option prices and the resulting option strategy will be a direct reflection of the view taken on traded quantities.

Part II: Calibration

CHAPTER 13

Calibrating the Hull–White extended Vasicek approach

In §7.3 we examine the pricing of contingent claims within the HW-extended Vasicek framework. The time t price of a European call option, with expiry time $T,\ t, T \in [0, T^*]$ and strike price X, on a zero coupon bond of maturity s is given by[1]:

$$C(t, T, s) = P(r, t, s)N(h) - XP(r, t, T)N(h - \sigma_p)$$

where

$$\begin{aligned} h &= \frac{1}{\sigma_p} \ln \frac{P(r, t, s)}{P(r, t, T)X} + \frac{\sigma_p}{2} \\ \sigma_p^2 &= (B(0, s) - B(0, T))^2 \int_t^T \left(\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}} \right)^2 d\tau \end{aligned}$$

In §7.5 we showed for the initial values of $B(t, T)$ i.e. $B(0, T)$, may be extracted from the observed term structure. Here we examine the approach to calibrating the HW extended Vasicek model to a given set of term structure data.

13.1. Using information from the observed term structure of interest rates and volatilities

Let $[0, T^*]$ be some trading interval. Data representing the term structure consists of the following:

$r(t)$	–	the instantaneous, continuously compounded short-term interest rate at time t,
$R(r, t, T)$	–	the interest rate term structure, that is continuously compounded rates for a series of maturity dates, $T,\ T \in [0, T^*]$,

[1]See equations (7.20)–(7.22) of Chapter 7.

$\sigma_r(r,t)$	–	instantaneous short-term interest rate volatility,
$\sigma_R(r,t,T)$	–	term structure of interest rate volatilities with maturities corresponding to those of the term structure of interest rates.

The price of a zero coupon bond takes the functional form:

$$P(r,t,T) = A(t,T)e^{-B(t,T)r(t)} \tag{13.1}$$

where $B(t,T)$ may be found from the time t term structure as (refer to equation (7.37) Chapter 7):

$$B(t,T) = \frac{R(r,t,T)\sigma_R(r,t,T)(T-t)}{r(t)\sigma_r(r,t)}$$

Letting $t = 0$ denote the current time, the term structure of $B(0,\cdot)$ is expressed in terms of the current volatility and interest rate term structures as:

$$B(0,T) = \frac{R(r,0,T)\sigma_R(r,0,T)T}{r(0)\sigma_r(r,0)} \tag{13.2}$$

Additionally, the initial term structure of interest rates $R(r,0,\cdot)$ may be used to determine the term structure of zero coupon bond prices or discount factors as:

$$P(r,0,T) = e^{-R(r,0,T)T} \tag{13.3}$$

This term structure of zero coupon bond prices, together with the term structure of $B(0,\cdot)$ allows the term structure of $A(0,\cdot)$ to be found as:

$$A(0,T) = P(r,0,T)e^{B(0,T)r(0)} \tag{13.4}$$

13.2. Call option on a coupon paying bond

The option pricing formula above may be used to obtain the value of a call option on a pure discount bond. However, these are not readily traded in all markets and one may be forced to use options on coupon paying bonds. Jamshidian [**31**] developed the following methodology allowing options on coupon paying bonds to be valued as portfolios of options on discount bonds.

Let $P_c(r,t,s)$ be the time t price of a coupon paying bond maturing at time s, $t,s \in [0,T^*]$. $P_c(r,t,s)$ consists of n payments c_i, $i = 1,\ldots,n$ at times s_i, $i = 1,\ldots,n$ where $s_i \in [t,s]$. The time t price of such a bond may be expressed as:

$$P_c(r,t,s) = \sum_{i=1}^{n} c_i P(r,t,s_i)$$

At option expiry time T, the bond has m payments remaining. Let r^* be the instantaneous continuously compounded short-term interest rate at time T, such that the price of the coupon paying bond equals the option strike price, i.e.:

$$P_c(r^*,T,s) = \sum_{i=1}^{m} c_i P(r^*,T,s_i) = X \tag{13.5}$$

The time T payoff of the European call option on such a coupon paying bond is:

$$\begin{aligned}
&\max\left[0, P_c(r,T,s) - X\right] \\
&= \max\left[0, \sum_{i=1}^{m} c_i P(r,T,s_i) - \sum_{i=1}^{m} c_i P(r^*,T,s_i)\right] \\
&= \sum_{i=1}^{m} c_i \max\left[0, P(r,T,s_i) - P(r^*,T,s_i)\right] \\
&= \sum_{i=1}^{m} c_i \max\left[0, P(r,T,s_i) - X_i\right]
\end{aligned}$$

where $X_i = P(r^*,T,s_i)$, $i = 1,\ldots,m$. Hence the payoff of the i^{th} option is:

$$\max\left[0, P(r,T,s_i) - X_i\right]$$

and the time t price of this option is:

$$C_i(t,T,s_i) = P(r,t,s_i)N(h_i) - X_i P(r,t,T)N(h_i - \sigma_{p_i})$$

with

$$h_i = \frac{1}{\sigma_{p_i}} \ln \frac{P(r,t,s_i)}{P(r,t,T)X_i} + \frac{\sigma_{p_i}}{2} \tag{13.6}$$

$$\sigma_{p_i}^2 = (B(0,s_i) - B(0,T))^2 \int_t^T \left(\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}} \right)^2 d\tau \tag{13.7}$$

13.2.1. Finding r^*. From equations (13.5) and (13.1) we may write the option strike price as:

$$X = \sum_{i=1}^{m} c_i P(r^*, T, s_i)$$

$$= \sum_{i=1}^{m} c_i A(T, s_i) e^{-B(T,s_i)r^*} \tag{13.8}$$

where (refer to equations (7.15) and (7.16) Chapter 7):

$$B(T, s_i) = \frac{B(0, s_i) - B(0,T)}{\frac{\partial B(0,T)}{\partial T}} \tag{13.9}$$

and

$$\hat{A}(T, s_i) = \hat{A}(0, s_i) - \hat{A}(0,T) - B(t, s_i)\frac{\partial \hat{A}(0,T)}{\partial T} - \frac{1}{2}\left[B(t, s_i)\frac{\partial B(0,T)}{\partial T}\right]^2 \int_0^T \left[\frac{\sigma(\tau)}{\frac{\partial B(0,\tau)}{\partial \tau}}\right]^2 d\tau \tag{13.10}$$

where

$$\hat{A}(T, s_i) = \ln A(T, s_i)$$

Since $B(T, s_i)$ and $\hat{A}(T, s_i)$ are fully specified by the initial term structure, we may apply a numerical search technique such as Newton–Raphson to solve for r^* such that (13.8) holds.

13.3. Using market data

Having considered the model formulation that allows us to incorporate observed term structure data into the pricing formulae, we examine how actual data is used in the calibration exercise.

For each day, the continuously compounded rate of interest and historical volatility are available for a discrete set of node points corresponding to terms to maturity, τ_i, $i = 1, \ldots, N$ where $\tau_1 = 1$ day.

- The interest rate with term to maturity 1 day and its corresponding historical volatility are taken as proxies for the instantaneous short-term interest rate $r(0)$ and its volatility $\sigma_r(r, 0)$.
- Making use of (13.2) we calculate $B(0,T)$ at each of the node points.
- Applying (13.3) we determine the time 0 discount bond prices, $P(r, 0, \cdot)$, with maturities at the nodes.
- These term structures of $B(0, \cdot)$ and $P(r, 0, \cdot)$ are applied in (13.4) to obtain the values of $A(0, \cdot)$ at the nodes.

An interpolation technique must be applied to the term structures of $A(0,\cdot)$ and $B(0,\cdot)$ so that values for any maturity term maybe extracted. Cubic spline interpolation was selected for the smoothness of curves it produces (see [**1**]).

13.4. Cubic spline interpolation

Cubic spline interpolation is a type of piecewise polynomial approximation that uses cubic polynomials between successive pairs of nodes [**13**]. Additionally, the constructed cubic piecewise interpolant is required to be twice continuously differentiable. This condition differentiates cubic spline interpolation from other types of cubic piecewise interpolation techniques, such as Cubic Hermite and Cubic Bessel interpolation (see [**20**]).

At each of the nodes across which the cubic splines are fitted, the following hold:

- The values of the fitted splines equal the values of the original function at the node points.
- The first and second derivatives of the fitted splines are continuous.

See Appendix for a detailed formulation of cubic spline interpolation.

As shown in (13.9) and (13.10), we require the derivatives of the initial term structures of $A(0,\cdot)$ and $B(0,\cdot)$. Although the construction of the cubic spline does not ensure the derivatives of the interpolant agree with the derivatives of the initial function [**13**], it does provide "acceptable approximations to derivatives" [**20**]. For this reason the derivatives of the fitted cubic splines are used as approximations to the required derivatives i.e. $\frac{\partial A(0,\cdot)}{\partial\cdot}$ and $\frac{\partial B(0,\cdot)}{\partial\cdot}$.

Now consider approximating the derivative at each node point x_j, by first fitting a quadratic polynomial to points x_{j-1}, x_j and x_{j+1} and then evaluating its derivative at x_j. Comparing these derivatives to those produced by taking the derivatives of the fitted cubic splines shows that the cubic splines produce much more extreme derivative values. In equations (13.7) and (13.10) the integral of the square of inverse of the derivative of $B(0,\cdot)$ is required. Here, the more extreme derivatives generated by the cubic splines give rise to inconsistencies and rather high integral values between certain node points. The effect is that of unreasonable future term structure shapes of $\hat{A}(t,\cdot)$ and unreasonably high volatility values for some of the sub-options constituting the coupon bond option. We investigate two ways of mitigating the magnitude of this effect:

13.4.1. Cubic Bessel interpolation.

Cubic Bessel interpolation (see [**20**]) uses the same basic methodology as Cubic Spline Interpolation, but places an additional constraint on the derivatives at the node points. The derivative at each node point x_j, is set equal to the derivative of the quadratic polynomial fitted to points x_{j-1}, x_j and x_{j+1}. For details of this procedure see the Appendix.

By forcing the derivatives at the nodes of the cubic interpolant to be equal to those of the quadratic polynomials, the derivatives between the nodes become less extreme and the integral of the square of the inverse of the derivative of the cubic polynomials becomes smoother. However, the integral between some of the short-term nodes is still rather large, again leading to overestimation of the volatility.

13.4.2. Interpolating the derivatives. The derivatives at the node points are assumed to equal those of the quadratic polynomials fitted as above. Cubic spline interpolation is then applied to these derivatives. This results in a smoothly interpolated derivative curve and hence smooth integral values. The integral required to evaluate (13.10) now becomes the integral of the square of the inverse of the cubic spline. The evaluation of such an integral is detailed in the Appendix.

13.5. Constant mean reversion and volatility parameters

The implementation of the HW-extended Vasicek model, as described above, represents a stochastic process of the short-term interest rate of the form[2]:

$$dr = [\theta(t) + a(t)(b - r)]\, dt + \sigma(t)\, dz \tag{13.11}$$

The three time-dependent parameters allow three characteristics of the initial term structure to be fitted (see [**29**]). The time-dependent drift parameter $\theta(t)$ allows the model to exactly match the initial interest rate term structure; the time-dependent short-term interest rate volatility $\sigma(t)$ defines the volatility of the short-term interest rate at times in the future; time-dependent reversion speed $a(t)$ specifies the relative volatilities of long and short-term interest rates, hence replicating the initial volatility term structure. However, fitting all three time-dependent parameters to current market data results in an over-parameterisation of the model, which may cause undesirable side effects [**29**]. Resulting future term structures of volatilities may take on implausible shapes, leading to mispricing of exotic options. For this reason implementations with constant reversion speed and short-term interest rate volatility a and σ respectively, are recommended. The volatility term structure will no longer be fitted exactly but only approximated. However, the model will display a stationary volatility term structure which allows more control over future values of model parameters and hence more accurate, robust pricing of exotic derivatives.

The resulting process of the short-term interest rate is:

[2]See equation (7.4).

$$dr = [\theta(t) + a(b - r)]\, dt + \sigma\, dz \tag{13.12}$$

with associated functional form of $B(t,T)$ reducing to[3]:

$$B(t,T) = \frac{1 - e^{-a(T-t)}}{a} \tag{13.13}$$

Hence the initial time $t = 0$ value is:

$$B(0,T) = \frac{1 - e^{-aT}}{a} \tag{13.14}$$

Since the time $t = 0$ term structure of interest rates is known we find the initial term structure of discount factors (zero coupon bond prices) $P(r, 0, \cdot)$, as shown in (13.3). Consequently the initial values of the $A(t,T)$ coefficient may be found as:

$$\begin{aligned} A(0,T) &= P(r,0,T)\, e^{\left(B(0,T) r(0)\right)} \\ &= P(r,0,T)\, e^{\left(\left(1-e^{-aT}\right) r(0)/a\right)} \end{aligned} \tag{13.15}$$

Since $P(r, 0, \cdot)$ and $r(0)$ are known, the exact values of $A(0, \cdot)$ are determined by the chosen value of parameter a. The initial discount bond prices can be exactly reproduced for any arbitrarily specified value of a. This does not say anything about the correctness of this parameter value. Incorporation of additional market data dependent on volatility parameters, such as interest rate options, is required to allow a correct choice of a.

As before, once the initial values $A(0, \cdot)$ and $B(0, \cdot)$ are known, any future values $A(t,T)$ and $B(t,T)$ may be calculated using (13.9) and (13.10). For constant a and σ these reduce to[4]:

$$B(T, s_i) = \frac{1 - e^{-a(s_i - T)}}{a} \tag{13.16}$$

$$\begin{aligned} \ln A(T, s_i) = \ln \frac{P(r,0,s_i)}{P(r,0,T)} &- \left(\frac{1 - e^{-a(s_i-T)}}{a}\right) \frac{\partial \ln P(r,0,T)}{\partial T} \\ &- \frac{\sigma^2}{4a}\left(1 - e^{-a(s_i-T)}\right)^2 \left(1 - e^{-2aT}\right) \end{aligned} \tag{13.17}$$

Similarly, the forward bond price volatility (13.7), required to value each of the m zero coupon bond options making up the coupon bond option, reduces to:

---

[3]See equation (7.25).

[4]These can easily be shown to be equivalent to the functional forms specified in (7.25), (7.26) Chapter 7 for constant a and σ.

$$\sigma_{p_i}^2 = \frac{\sigma^2}{2a^3}\left(1 - e^{-a(s_i - T)}\right)^2\left(1 - e^{-2a(T-t)}\right) \tag{13.18}$$

Calibration of the model to observable market prices involves retrieving values of σ and a such that these market prices may be recovered from the model.

13.6. Flat volatility term structure

The functional form of $B(t,T)$ considered in §13.5 does not allow $a = 0$. This is the case of zero mean reversion. Assuming a constant, flat volatility structure and consequently a zero mean reversion parameter, the equation of the short-term interest rate process becomes:

$$dr = \theta(t)\,dt + \sigma\,dz \tag{13.19}$$

Consider the relationship of $B(0,T)$ to the initial interest rate and volatility term structure as shown in (13.2):

$$B(0,T) = \frac{R(r,0,T)\sigma_R(r,0,T)T}{r(0)\sigma_r(r,0)} \tag{13.20}$$

Allowing a constant volatility structure implies:

$$\sigma_r(r,t) = \sigma_R(r,t,T) = \sigma \qquad \forall\, t,T \in [0,T^*]\,\text{and}\, t < T$$

Hence (13.20) reduces to:

$$B(0,T) = \frac{R(r,0,T)}{r(0)}T \tag{13.21}$$

To show that this does in fact imply a zero reversion speed parameter, consider equation (7.17) in Chapter 7:

$$\begin{aligned} a(t) &= -\left(\frac{\partial^2 B(0,t)}{\partial t^2}\right)\Big/\left(\frac{\partial B(0,t)}{\partial t}\right) \\ &= -\frac{0}{\frac{R(r,0,T)}{r(0)}} = 0 \qquad \forall\, t \in [0,T^*] \end{aligned}$$

Making use of (13.1) to determine the functional form of $A(0,\cdot)$ we have:

$$\begin{aligned} P(r,0,T) &= A(0,T)e^{-B(0,T)r(0)} \\ &= A(0,T)e^{-R(r,0,T)T} \end{aligned}$$

However, from (13.3)

$$P(r,0,T) = e^{-R(r,0,T)T} \tag{13.22}$$

Hence $A(0,T) = 1$ for all $T \in [0,T^*]$.

The calibration of the model now requires the fitting of a single volatility parameter σ such that market prices of traded securities are recovered as closely as possible.

13.7. Calibration methodology

13.7.1. Algorithm for constant mean reversion and volatility parameters. The following is an algorithm of the methodology to calibrate a reversion speed a and associated volatility parameter σ which give rise to the smallest mispricing.

```
For valuation_date 1 to n
  a = initial_value
  While a is in the acceptable interval
    For expiry 1 to m
      initialise σ_1 and σ_2
      get Black_premium for this expiry and valuation date
      P_1 = HW_option_premium (σ_1) - Black_premium
      P_2 = HW_option_premium (σ_2) - Black_premium
      While |P_1| > 1 * 10^-5 and |P_2| > 1 * 10^-5
        σ_mid = (σ_1 + σ_2)/2
        P_mid = HW_option_premium (σ_mid) - Black_premium
        If P_1 * P_mid < 0 Then
          σ_2 = σ_mid
        ElseIf P_2 * P_mid < 0 Then
          σ_1 = σ_mid
        End If
      Loop
      σ = σ_mid
    Next expiry
    σ_ave = Average(σ for expiry 1 to m)
    ave_mispricing = Average((HW_option_premium (σ_ave)
                        − Black_premium)/Black_premium;
                        for expiry 1 to m)
    If ave_mispricing < optimal_mispricing then a_optimal = a
    update a
  Loop
Next valuation_date
```

For a given value of parameter a, determine the corresponding values of σ_i for expiry date $i = 1, \ldots, m$. Since a and σ are parameters associated with the term structure, specifically the reversion speed and volatility of the short-term interest rate, they tell us what the price of the bond option implies about the characteristics of the short-term interest rate. Ideally, we would like the σ_is to be the same for all maturities $i = 1, \ldots, m$. This would imply all maturity bond options are priced consistently by our model of the short-term interest rate. However, in practice these σ_is may differ quite substantially. To determine an optimal value of a and corresponding σ, take the arithmetic average of σ_i, $i = 1, \ldots, m$ to be the proxy for σ. Pricing each bond option using the value of a and corresponding σ allows us to determine the degree of mispricing to market-observed option premia. The optimal value of a and corresponding σ (calculated using the averaging described above) are determined as those resulting in the smallest mispricing across the bond option maturities.

The methodology used in procedure *HW_option_premium* to calculate the HW option premium is as described in §13.5. To calculate the strike prices and hence premia of the sub-options comprising the coupon bond options we require $A(T, s_i)$ where T is the option expiry date and s_i is a coupon payment date. From equation (13.17) we require $P(r, 0, s_i)$, $P(r, 0, T)$ and $\frac{\partial \ln P(r,0,T)}{\partial T}$ in order to evaluate $A(T, s_i)$. Since T and s_i will not necessarily fall on the available node points, we must interpolate the $P(r, 0, \cdot)$ and $\frac{\partial \ln P(r,0,\cdot)}{\partial \cdot}$ curves to obtain values for the correct maturity dates.

Given the initial interest rate term structure $R(r, 0, \cdot)$, make use of (13.3) to obtain the term structure of $P(r, 0, \cdot)$. We then apply the cubic spline interpolation to retrieve discount bond prices for any maturity date.

In order to evaluate $\frac{\partial \ln P(r,0,\cdot)}{\partial \cdot}$, we first calculate $\ln P(r, 0, \cdot)$ at each node point and then apply Lagrange interpolation to fit quadratic polynomials and hence evaluate derivatives at each node[5]. Applying cubic spline interpolation to these derivative values allows us to retrieve $\frac{\partial \ln P(r,0,\cdot)}{\partial \cdot}$ for any maturity date.

13.7.2. Algorithm for a flat volatility term structure. The following algorithm describes the methodology to calibrate a volatility parameter σ which gives rise to the smallest mispricing. This algorithm is essentially a subset of the previous algorithm since the methodology corresponds to $a = 0$ and hence only an optimal σ needs to be found.

[5] This is the methodology described in the final paragraph of §13.4.1.

```
For valuation_date 1 to n
  For expiry 1 to m
    initialise σ1 and σ2
    get Black_premium for this expiry and valuation date
    P1 = HW_option_premium (σ1) - Black_premium
    P2 = HW_option_premium (σ2) - Black_premium
    While |P1| > 1 * 10^-5 and |P2| > 1 * 10^-5
      σmid = (σ1 + σ2)/2
      Pmid = HW_option_premium (σmid) - Black_premium
      If P1 * Pmid < 0 Then
        σ2 = σmid
      ElseIf P2 * Pmid < 0 Then
        σ1 = σmid
      End If
    Loop
    σ = σmid
  Next expiry
Next valuation_date
```

To determine the strikes of the sub-options and hence their premia we require values of $B(T, s_i)$ and $A(T, s_i)$, where T is the option expiry date and s_i the coupon payment date.

To determine $B(T, s_i)$ we must determine[6] $B(0, \cdot)$ and $\frac{\partial B(0,\cdot)}{\partial \cdot}$. Given the initial interest rate term structure $R(r, 0, \cdot)$, make use of (13.21) to determine the values of $B(0, \cdot)$ at all the node points. Applying cubic spline interpolation allows us to retrieve $B(0, \cdot)$ for any maturity date. To determine $\frac{\partial B(0,\cdot)}{\partial \cdot}$, we fit a quadratic Lagrange polynomial at each node point and evaluate its first derivative. Then applying cubic spline interpolation to these derivatives allows us to retrieve $\frac{\partial B(0,\cdot)}{\partial \cdot}$ for any maturity.

To determine $A(T, s_i)$ we also require $\int_0^T \left(1/\frac{\partial B(0,\tau)}{\partial \tau}\right)^2 d\tau$. This is found by applying cubic spline interpolation to $\frac{\partial B(0,\cdot)}{\partial \cdot}$ and then evaluating this integral. This involves the integral of the square of a cubic polynomial.

[6]See equation (13.9).

Appendix

Cubic Spline Interpolation

Given a continuous function f on interval $[a, b]$ with nodes $x_i,\ i = 0, \ldots, n$ such that $a = x_0 < x_1 < \cdots < x_n = b$, a cubic spline interpolant S, is a function such that:

(1) The cubic polynomial on the subinterval $[x_j, x_{j+1}]$, $j = 0, 1, \ldots, n-1$ is denoted $S(x) = S_j(x)$;
(2) $S(x_j) = f(x_j)$ for all $j = 0, 1, \ldots, n$;
(3) $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$ for all $j = 0, 1, \ldots, n-2$;
(4) $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$ for all $j = 0, 1, \ldots, n-2$;
(5) $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$ for all $j = 0, 1, \ldots, n-2$;
(6) Since we have no conclusive information about the second derivative of function f at the boundary nodes, we fit a free or natural cubic spline by imposing the condition:

$$S''(x_0) = S''(x_n) = 0$$

Here, the cubic polynomials take the form:

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3 \qquad \text{for all } j = 0, 1, \ldots, n-2 \tag{A.1.}$$

Applying conditions (1)–(5) to the set of polynomials in (A.1.), we arrive at a linear system of equations:

$$h_{j-1}c_{j-1} + 2(h_{j-1} - h_j)c_j + h_j c_{j+1} = \frac{3}{h_j}(a_{j+1} - a_j) - \frac{3}{h_{j-1}}(a_j - a_{j-1})$$

where

$$h_j = x_{j+1} - x_j$$

and condition (6) gives:

$$c_0 = c_n = 0$$

The above equations form a tridiagonal linear system, which may be solved for a unique solution (see [**13**]).

Cubic Bessel Interpolation

Using Lagrange interpolation, fit a quadratic polynomial $p(x)$, to node points x_{j-1}, x_j, x_{j+1}. This takes the form

$$
\begin{aligned}
p(x) =& f(x_{j-1})\frac{(x-x_j)(x-x_{j+1})}{(x_{j-1}-x_j)(x_{j-1}-x_{j+1})} + f(x_j)\frac{(x-x_{j-1})(x-x_{j+1})}{(x_j-x_{j-1})(x_j-x_{j+1})} \\
&+ f(x_{j+1})\frac{(x-x_{j-1})(x-x_j)}{(x_{j+1}-x_{j-1})(x_{j+1}-x_j)}
\end{aligned}
$$

hence

$$
\begin{aligned}
p'(x) =& f(x_{j-1})\frac{2x-x_j-x_{j+1}}{(x_{j-1}-x_j)(x_{j-1}-x_{j+1})} + f(x_j)\frac{2x-x_{j-1}-x_{j+1}}{(x_j-x_{j-1})(x_j-x_{j+1})} \\
&+ f(x_{j+1})\frac{2x-x_{j-1}-x_j}{(x_{j+1}-x_{j-1})(x_{j+1}-x_j)}
\end{aligned}
$$

and so evaluating this derivative at node x_j we have:

$$
\begin{aligned}
p'(x_j) =& f(x_{j-1})\frac{x_j-x_{j+1}}{(x_{j-1}-x_j)(x_{j-1}-x_{j+1})} + f(x_j)\frac{2x_j-x_{j-1}-x_{j+1}}{(x_j-x_{j-1})(x_j-x_{j+1})} \\
&+ f(x_{j+1})\frac{x_j-x_{j-1}}{(x_{j+1}-x_{j-1})(x_{j+1}-x_j)}
\end{aligned}
$$

Hence we find the polynomial S_j, $j = 0, 1, \ldots, n-2$ such that:

$$
\begin{aligned}
&S_j(x_j) = f(x_j) \qquad S_{j+1}(x_{j+1}) = f(x_{j+1}) \\
&S'_j(x_j) = p'(x_j) \qquad S'_{j+1}(x_{j+1}) = p'(x_{j+1}) \qquad \text{for } j = 0, 1, \ldots, n-2
\end{aligned}
$$

This gives rise to coefficient values defined by (see [**20**]):

$$
\begin{aligned}
a_j &= f(x_j) \\
b_j &= p'(x_j) \\
c_j &= \frac{f(x_{j+1})-f(x_j)}{(x_{j+1}-x_j)^2} - \frac{p'(x_j)}{x_{j+1}-x_j} - d_j(x_{j+1}-x_j) \\
d_j &= \frac{p'(x_j)+p'(x_{j+1})}{(x_{j+1}-x_j)^2} - 2\frac{f(x_{j+1})-f(x_j)}{(x_{j+1}-x_j)^3}
\end{aligned}
$$

Integrating the inverse of the square of a cubic polynomial

The required integral has the form:

(A.2.) $$\int_m^n \left(\frac{1}{d_j(x-x_j)^3+c_j(x-x_j)^2+b_j(x-x_j)+a_j}\right)^2 dx$$

First, consider a cubic polynomial of the form:

$$dx^3 + cx^2 + bx + a = 0$$

then according to the methodology detailed in [**42**] define the following variables:

(A.3.a) $$x_N = -\frac{c}{3d}$$

(A.3.b) $$y_N = dx_N^3 + cx_N^2 + bx_N + a$$

(A.3.c) $$\delta^2 = \frac{c^2 - 3db}{9d^2}$$

(A.3.d) $$h^2 = 4d^2\delta^6$$

Now, three cases must be considered:

(1) $y_N^2 - h^2 > 0$;
(2) $y_N^2 - h^2 = 0$;
(3) $y_N^2 - h^2 < 0$.

$y_N^2 - h^2 > 0$. Here, the cubic polynomial has only one real root, given by:

$$x_1 = x_N + \sqrt[3]{\frac{1}{2d}\left(-y_N + \sqrt{y_N^2 - h^2}\right)} + \sqrt[3]{\frac{1}{2d}\left(-y_N - \sqrt{y_N^2 - h^2}\right)}$$

Hence:

$$\begin{aligned}
&dx^3 + cx^2 + bx + a \\
&= \big(x - x_1\big)\big(dx^2 + \big(c + dx_1\big)x + b + cx_1 + dx_1^2\big) \\
&= d\big(x - x_1\big)\Big(x^2 + \big(\frac{c}{d} + x_1\big)x + \frac{b}{d} + \frac{c}{d}x_1 + x_1^2\Big) \\
&= d\big(x - x_1\big)\big(x^2 + \mu x + \rho\big)
\end{aligned}$$

where $\mu = \frac{c}{d} + x_1$ and $\rho = \frac{b}{d} + \frac{c}{d}x_1 + x_1^2$.

Applying this to (A.2.) the integral takes the form:

$$\begin{aligned}
&\left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{1}{(x - x_j - x_1)((x - x_j)^2 + \mu(x - x_j) + \rho)}\right)^2 dx \\
&= \left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{G}{x - x_j - x_1} + \frac{I - G(x - x_j)}{(x - x_j)^2 + \mu(x - x_j) + \rho}\right)^2 dx
\end{aligned}$$

where:

$$G = \frac{1}{x_1^2 + \mu x_1 + \rho} \qquad I = -G(x_1 + \mu)$$

Further:

$$\begin{aligned}
&\left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{G^2}{(x-x_j-x_1)^2} + \frac{\big(I-G(x-x_j)\big)^2}{\big((x-x_j)^2+\mu(x-x_j)+\rho\big)^2}\right. \\
&\quad \left. + \frac{2G\big(I-G(x-x_j)\big)}{\big(x-x_j-x_1\big)\big((x-x_j)^2+\mu(x-x_j)+\rho\big)}\right) dx \\
&= \left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{G^2}{(x-x_j-x_1)^2} + \frac{G^2}{(x-x_j)^2+\mu(x-x_j)+\rho}\right. \\
&\quad + \frac{L(x-x_j)+M}{\big((x-x_j)^2+\mu(x-x_j)+\rho\big)^2} - \frac{P}{x-x_j-x_1} \\
&\quad \left. + \frac{P(x-x_j)+Q}{(x-x_j)^2+\mu(x-x_j)+\rho}\right) dx
\end{aligned}$$

where:

$$L = -2GI - G^2\mu \qquad M = I^2 - G^2\rho$$

$$P = \frac{2G^2(2x_1+\mu)}{x_1^2+\mu x_1+\rho} \qquad Q = \frac{2G^2\big((\mu+x_1)^2-\rho\big)}{x_1^2+\mu x_1+\rho}$$

and so:

$$\begin{aligned}
&\left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{G^2}{(x-x_j-x_1)^2} - \frac{P}{x-x_j-x_1}\right. \\
&\quad + \left(\frac{P}{2}\right)\frac{2(x-x_j)+\mu}{(x-x_j)^2+\mu(x-x_j)+\rho} + \frac{Q+G^2-\frac{1}{2}P\mu}{(x-x_j)^2+\mu(x-x_j)+\rho} \\
&\quad \left. + \left(\frac{L}{2}\right)\frac{2(x-x_j)+\mu}{\big((x-x_j)^2+\mu(x-x_j)+\rho\big)^2} + \frac{M-\frac{1}{2}L\mu}{\big((x-x_j)^2+\mu(x-x_j)+\rho\big)^2}\right) dx \\
&= \left(\frac{1}{d_j}\right)^2 \left[\frac{-G^2}{x-x_j-x_1} - P\ln|x-x_j-x_1| + \frac{P}{2}\ln\left|(x-x_j)^2+\mu(x-x_j)+\rho\right|\right. \\
&\quad - \frac{\frac{1}{2}L}{(x-x_j)^2+\mu(x-x_j)+\rho} + \frac{M-L\eta}{2\phi^2}\left(\frac{x-x_j+\eta}{(x-x_j)^2+\mu(x-x_j)+\rho}\right) \\
&\quad \left. + \left(\frac{Q+G^2-P\eta}{\phi} + \frac{M-L\eta}{2\phi^3}\right)\arctan\left(\frac{x-x_j+\eta}{\phi}\right)\right]_{x=m}^{x=n}
\end{aligned}$$

where $\eta = \frac{1}{2}\mu$ and $\phi = \sqrt{\rho-\eta^2}$.

$y_N^2 - h^2 = 0$**.** For this case there are three real roots, with a repeated root. Again from the methodology in [**42**], the three roots are:

$$x_1 = x_N - 2\sqrt[3]{\frac{y_N}{2d}}$$
$$x_2 = x_3 = x_N - \sqrt[3]{\frac{y_N}{2d}}$$

Substituting into (A.2.) we have:

$$\begin{aligned}
&\left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{1}{(x-x_j-x_1)(x-x_j-x_2)^2}\right)^2 dx \\
&= \left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{A}{x-x_j-x_1} - \frac{A}{x-x_j-x_2} + \frac{C}{(x-x_j-x_2)^2}\right)^2 dx \\
&= \left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{A^2}{(x-x_j-x_1)^2} + \frac{A^2+H}{(x-x_j-x_2)^2} - \frac{2AC}{(x-x_j-x_2)^3}\right. \\
&\quad \left. + \frac{C^2}{(x-x_j-x_2)^4} + \frac{F-D}{x-x_j-x_1} - \frac{F-D}{x-x_j-x_2}\right) dx \\
&= \left(\frac{1}{d_j}\right)^2 \left[-\frac{A^2}{x-x_j-x_1} - \frac{A^2+H}{x-x_j-x_2} + \frac{AC}{(x-x_j-x_2)^2}\right. \\
&\quad \left. - \frac{\frac{1}{3}C^2}{(x-x_j-x_2)^3} + (F-D)\ln\left|\frac{x-x_j-x_1}{x-x_j-x_2}\right|\right]_{x=m}^{x=n}
\end{aligned}$$

where:

$$A = \frac{1}{(x_1-x_2)^2} \qquad C = \frac{-1}{x_1-x_2}$$
$$D = \frac{2A^2}{x_1-x_2} \qquad F = \frac{2AC}{(x_2-x_1)^2} \qquad H = \frac{2AC}{x_2-x_1}$$

$y_N^2 - h^2 < 0$. Here we have three distinct real roots. These roots are:

$$x_1 = x_N + 2\delta\cos\theta$$
$$x_2 = x_N + 2\delta\cos\left(\theta + \frac{2\pi}{3}\right)$$
$$x_3 = x_N + 2\delta\cos\left(\theta + \frac{4\pi}{3}\right)$$

where[7]:

$$\theta = \tfrac{1}{3}\arccos\left(\frac{-y_N}{h}\right)$$

[7]As defined in (A.3.d), $h = 2d\delta^3$.

Again, substituting into (A.2.), we evaluate the integral as:

$$
\begin{aligned}
&\left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{1}{(x-x_j-x_1)(x-x_j-x_2)(x-x_j-x_3)}\right)^2 dx \\
&= \left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{A}{x-x_j-x_1} + \frac{B}{x-x_j-x_2} + \frac{C}{x-x_j-x_3}\right)^2 dx \\
&= \left(\frac{1}{d_j}\right)^2 \int_m^n \left(\frac{A^2}{(x-x_j-x_1)^2} + \frac{B^2}{(x-x_j-x_2)^2} + \frac{C^2}{(x-x_j-x_3)^2}\right. \\
&\quad \left. + \frac{D+F}{x-x_j-x_1} + \frac{H-D}{x-x_j-x_2} - \frac{H+F}{x-x_j-x_3}\right) dx \\
&= \left(\frac{1}{d_j}\right)^2 \left[-\frac{A^2}{x-x_j-x_1} - \frac{B^2}{x-x_j-x_2} - \frac{C^2}{x-x_j-x_3}\right. \\
&\quad + (D+F)\ln|x-x_j-x_1| + (H-D)\ln|x-x_j-x_2| \\
&\quad \left. - (H+F)\ln|x-x_j-x_3| \right]_{x=m}^{x=n}
\end{aligned}
$$

where:

$$
\begin{aligned}
A &= \frac{x_2-x_3}{x_2^2(x_3-x_1) - x_3^2(x_2-x_1) - x_1^2(x_3-x_2)} \\
B &= \frac{x_3-x_1}{x_2^2(x_3-x_1) - x_3^2(x_2-x_1) - x_1^2(x_3-x_2)} \\
C &= \frac{x_1-x_2}{x_2^2(x_3-x_1) - x_3^2(x_2-x_1) - x_1^2(x_3-x_2)}
\end{aligned}
$$

and

$$
D = \frac{2AB}{x_1-x_2} \qquad F = \frac{2AC}{x_1-x_3} \qquad H = \frac{2BC}{x_2-x_3}
$$

CHAPTER 14

Calibrating the Black, Derman and Toy discrete time model

In Chapter 8 we examined the formulation of the BDT model within a (discrete time) binomial lattice as well as its continuous time equivalent. The short-term interest rate process takes the form[1]:

$$r(t) = u(t) \exp\left(\sigma(t) z(t)\right)$$

where $r(t)$, $u(t)$, $\sigma(t)$ and $z(t)$ are the time t values of the short-term interest rate, the median of the (lognormal) short-term interest rate distribution, the short-term interest rate volatility and a standard Brownian motion. The continuous time equivalent of the BDT model is determined to be:

$$d \ln r(t) = \left[\frac{\partial \ln u(t)}{\partial t} - \frac{\partial \ln \sigma(t)}{\partial t}\left(\ln u(t) - \ln r(t)\right)\right] dt + \sigma(t) dz$$

However, this formulation of the model is not analytically tractable (this is a characteristic of all lognormal models) it must be implemented by means of a binomial tree. The tree is constructed so as to approximate the above stochastic differential equation of the short-term interest rate. The short-term interest rate at each node of the tree is determined so as to be consistent with observed market prices of bonds.

14.1. Initial Term Structure

Following the notation and methodology of Chapter 8 §8.5.1, the initial term structure of interest rates and interest rate volatilities is described by the following variables:

$P(i)$	–	time 0 price of a discount bond maturing at time $i\Delta t$,
$R(i)$	–	time 0 (continuously compounded) yield on a discount bond maturing at time $i\Delta t$,
$\sigma_R(i)$	–	time 0 volatility of yield $R(i)$.

[1]This relationship is derived in §8.4 of Chapter 8.

14.2. Calibrating to interest rate term structure only

As discussed in Chapter 8, the rate of mean reversion is a function of the short-term interest rate volatility. This implies that, by specifying the functional form of the time-dependent short-term interest rate volatility, we are simultaneously specifying the shape of the term structure of volatilities and vice versa. This may not be a desirable property, since future volatility term structures may be distorted, taking on unreasonable characteristics. For this reason many practitioners allow for a constant short-term interest rate volatility parameter when using the BDT model. This implies that the mean reversion speed is zero and the model is calibrated to the initial interest rate term structure only. A constant volatility parameter may also prove optimal in markets where a reliable term structure of implied interest rate volatilities is not available and historical volatilities are seen to be poor proxies for implied volatilities.

Letting $\sigma(t) = \sigma$ where σ is constant, the continuous time short rate dynamics reduce to:

$$d\ln r(t) = \frac{\partial \ln u(t)}{\partial t}\, dt + \sigma dz \tag{14.1}$$

and from (8.17) in Chapter 8, the discrete time representation of the short-term interest rate becomes:

$$r_{i,j} = u(i) \exp\left(\sigma j \sqrt{\Delta t}\right) \tag{14.2}$$

Hence calibration of the binomial tree to market data reduces to finding $u(i)$ at each time step. Given a value of σ we may calibrate the binomial tree such that market-observed discount bond prices are retrieved. Since the bond price is independent of σ, this gives no clue as to the correctness or otherwise of the chosen σ value. To ascertain the correctness of the σ parameter we need to make use of another set of security prices which depend on interest rate volatility (i.e. σ).

14.3. Forward Induction: making use of state prices

The required calibration procedure is a simpler version of that outlined in Chapter 8, §8.5. Again we make use of forward induction and state or Arrow–Debreu prices. Define $Q_{i,j}$ as the time 0 value of security paying 1 if node (i, j) is reached and 0 otherwise. By definition $Q_{0,0} = 1$.

These Arrow–Debreu prices may be seen as discounted probabilities and hence may be used to specify the value of any instrument. Specifically, the time 0 price of a discount bond maturing at time $(i + 1)\Delta t$ may be written as:

$$P(i+1) = \sum_j Q_{i,j} d_{i,j} \tag{14.3}$$

where $d_{i,j}$ is the one period discount factor at node (i,j) expressed as:

$$d_{i,j} = \frac{1}{1 + r_{i,j}\Delta t} \tag{14.4}$$

At each node (i,j) of the tree, the state prices are determined as functions of the state prices at time $(i-1)$ using:

$$Q_{i,i} = \tfrac{1}{2}\, Q_{i-1,i-1} d_{i-1,i-1} \tag{14.5a}$$

$$Q_{i,j} = \tfrac{1}{2}\, Q_{i-1,j-1} d_{i-1,j-1} + \tfrac{1}{2}\, Q_{i-1,j+1} d_{i-1,j+1} \qquad \forall\, j \neq i, -i \tag{14.5b}$$

$$Q_{i,-i} = \tfrac{1}{2}\, Q_{i-1,-i+1} d_{i-1,-i+1} \tag{14.5c}$$

Initial step: Initialise variables at time $i = 0$ as follows:

$$\begin{aligned} Q_{0,0} &= 1, \\ u(0) = r_{0,0} &= \frac{e^{(R(1)\Delta t)} - 1}{\Delta t} \qquad \text{and} \\ d_{0,0} &= \frac{1}{1 + r_{0,0}\Delta t} \end{aligned}$$

Now, for $i > 0$, we assume the following are known for all states j at time step $(i-1)$: $Q_{i-1,j}$, $u(i-1)$, $r_{i-1,j}$ and $d_{i-1,j}$. Hence, the values of $Q_{i,j}$, $u(i)$, $r_{i,j}$ and $d_{i,j}$ may be found for all states j at time i as follows:

Step 1: Make use of (14.5) to generate $Q_{i,j}$ as follows:

$$\begin{aligned} Q_{i,i} &= \tfrac{1}{2}\, Q_{i-1,i-1} d_{i-1,i-1} \\ Q_{i,-i} &= \tfrac{1}{2}\, Q_{i-1,-i+1} d_{i-1,-i+1} \\ &\text{For } j = -\,i+2 \text{ to } i-2 \text{ Step } 2 \\ Q_{i,j} &= \tfrac{1}{2}\, Q_{i-1,j-1} d_{i-1,j-1} + \tfrac{1}{2}\, Q_{i-1,j+1} d_{i-1,j+1} \end{aligned}$$

Step 2: A numerical search technique such as Newton–Raphson (e.g. see [**13**]) is used to find $u(i)$ such that the following is satisfied[2]:

[2]Here the summation is over all states j at time step i, hence $j = -i, -i+2, \ldots, i-2, i$.

$$
\begin{aligned}
P(i+1) &= \sum_j Q_{i,j} d_{i,j} \\
&= \sum_j Q_{i,j} \frac{1}{1 + u(i) \exp\left(\sigma j \sqrt{\Delta t}\right) \Delta t}
\end{aligned}
$$

Step 3: Using the $u(i)$s calculated in Step 2, the short-term interest rates $r_{i,j}$, and corresponding discount factors $d_{i,j}$, are updated for each state j at time step i as:

$$
\begin{aligned}
r_{i,j} &= u(i) \exp\left(\sigma j \sqrt{\Delta t}\right) \\
d_{i,j} &= \frac{1}{1 + r_{i,j} \Delta t}
\end{aligned}
$$

Steps 1–3 are repeated for all $i = 1, \ldots, N$ where $N\Delta t$ is the longest maturity discount bond.

14.4. Pricing contingent claims – Backward Induction

Once the short-term interest rate tree has been constructed such that the short-term interest rate and associated discount factor are known at each node of the tree, any interest rate contingent claim may be priced by a simple backward induction procedure.

14.4.1. Pricing Coupon Bonds. Consider a coupon bond maturing at time[3] $T = N_T \Delta t$ and paying coupons at discrete time steps $\{t_1, t_2, \ldots, t_m\}$ where m is the number of coupons due until maturity. If c is the amount payable at each coupon time and the last coupon payment coincides with bond maturity, i.e. $t_m = N_T$, the maturity value of the bond may be initialised at time $i = N_T$ as:

$$
P^c_{N_T,j} = 1 + c
$$

where $P^c_{i,j}$ is the value of the coupon paying bond at node (i, j).

Now for $i < N_T$, the value of the coupon bond is equal to the discounted expected value of the bond at the next time step $(i+1)$. Since the risk-neutral probability associated with each branch of the binomial tree is $\frac{1}{2}$, the value of the coupon bond for all $i < N_T$ is determined as[4] follows.

[3] Hence, this bond may be seen to mature N_T time steps from initial time $i = 0$.

[4] Here j represents each node at time step i, hence $j = -i, -i+2, \ldots, i-2, i$.

If $i \in \{t_1, t_2, \ldots, t_m\}$ then

$$P^c_{i,j} = \tfrac{1}{2}\, d_{i,j} \Big(P^c_{i+1,j+1} + P^c_{i+1,j-1} \Big) + c$$

else

$$P^c_{i,j} = \tfrac{1}{2}\, d_{i,j} \Big(P^c_{i+1,j+1} + P^c_{i+1,j-1} \Big)$$

14.4.2. Pricing European Options on Coupon Bonds. Once we have determined the value of the coupon paying bond at each node of the binomial tree, we may price claims contingent on this coupon bond. Consider a European call option on the above coupon paying bond with:

- expiry date $s = N_s \Delta t$,
- strike price X.

Knowing the value of the coupon bond in all states at option expiry time $i = N_s$, we determine the option payoff as:

$$C_{N_s,j} = \max\{0, P^c_{N_s,j} - X\} \qquad \text{for each } j = -N_s, \ldots, N_s \tag{14.6}$$

where $C_{i,j}$ is the value of the European call option at node (i, j). For each $i < N_s$ the value of the European call is determined as the discounted expectation of option values at time $(i + 1)$; hence:

$$C_{i,j} = \tfrac{1}{2}\, d_{i,j} \Big(C_{i+1,j+1} + C_{i+1,j-1} \Big)$$

More directly, utilise the discounted probabilities or state prices $Q_{i,j}$ to determine the time $i = 0$ European option value directly from the expiry condition (14.6) as:

$$C_{0,0} = \sum_j Q_{N_s,j} \max\{0, P^c_{N_s,j} - X\}$$

14.5. Calibration methodology for a constant volatility parameter

14.5.1. Algorithm. The following algorithm illustrates the basic methodology for calibrating volatility parameter σ such that traded options are priced with smallest error.

14.5.1.1. *Main procedure.* Main procedure within which the optimisation is performed.

```
For valuation_date 1 to n
   For expiry_date 1 to m
      initialise σ1 and σ2
      get Black_premium for this expiry and valuation date

      Calibrate_Rate_Tree (σ1)
      Evaluate_Bond_Tree
      P1 = BDT_option_premium (σ1) - Black_premium

      Calibrate_Rate_Tree (σ2)
      Evaluate_Bond_Tree
      P2 = BDT_option_premium (σ2) - Black_premium

      While |P1| > 1 * 10^-5 and |P2| > 1 * 10^-5
         σ_mid = (σ1 + σ2)/2
         Calibrate_Rate_Tree (σ_mid)
         Evaluate_Bond_Tree
         P_mid = BDT_option_premium (σ_mid) - Black_premium
         If P1 * P_mid < 0 Then
            σ2 = σ_mid
         ElseIf P2 * P_mid < 0 Then
            σ1 = σ_mid
         End If
      Loop
      σ = σ_mid
   Next expiry_date
Next valuation_date
```

14.5.1.2. *Sub-procedures.* The procedure *Calibrate_Rate_Tree* builds the binomial tree of short-term interest rates for the specified volatility parameter. The basic procedure followed is as described in §14.3. Once the tree of short-term interest rates has been constructed, the coupon bond is priced by procedure *Evaluate_Bond_Tree.* Due to the long maturity dates of the coupon bonds it is not practical to construct trees with daily time steps. Using a larger time step, say 30 days, introduces some complications. Since bond maturity and coupon payments may fall between node points, we need to interpolate these values to maintain accuracy. Below is the algorithm used to price the coupon bond for time steps greater than one day. We let N_b be the final time step of the tree. This will be either equal to or just greater than, the maturity of the bond underlying the option we wish to price.

```
If bond maturity lies between time N_b − 1 and N_b then
   offset = [(Maturity_Date − Valuation_Date)/Time_Step − (N_b − 1)]
         ∗Time_Step/365
   For each state j = −(N_b − 1) to (N_b − 1) Step 2
      B(N_b − 1, j) = (1 + Coupon)/(1 + r(N_b − 1, j) ∗ offset)
   Next state j
End if

For each time i = (N_b − 2) to 0 Step −1
   If a coupon lies between times i and i + 1 then
      offset = [Coupon_Term/Time_Step − i] ∗ Time_Step/365
      For state j = −i to i Step 2
         Coupon_Value = Coupon/(1 + r_{i,j} ∗ offset)
         B_{i,j} = 1/2 d_{i,j} [B_{i+1,j+1} + B_{i+1.j−1}] + Coupon_Value
      Next state j
   Else
      For state j = −i to i Step 2
         B_{i,j} = 1/2 d_{i,j} [B_{i+1,j+1} + B_{i+1.j−1}]
      Next state j
   End If
Next time i
```

A similar problem is encountered when pricing the coupon bond option in the procedure *BDT_option_premium*. The option expiry date may fall between two time nodes. Further complications may arise from the $T + \alpha$ settlement convention for bonds[5]. The strike refers to a bond price at option expiry + α. To calculate the terminal option value we must discount the strike price to option expiry. Option expiry and (option expiry+α) may occur between the same pair of time nodes, on either side of a time node or either one, but not both, may occur at a time node (unless time step is α days). These possibilities must all be taken into account to allow for correct interpolation of the strike price and bond price. This algorithm is outlined below.

[5]This represents time T plus α business days.

Interpolate to determine bond price and option strike price at option expiry.

```
If settlement date of option expiry lies between time i and i+1 then
   offset = [Settlement_Term/Time_Step − i] ∗ Time_Step/365
   For state j = −i to i Step 2
      Strike_{i,j} = Strike_AIP/(1 + r_{i,j} ∗ offset)
   Next state j
Else
   For state j = −i to i Step 2
      Strike_{i,j} = Strike_AIP
   Next state j
End If

If option expiry date lies between time i and i+1 then
   For state j = −i to i Step 2
      Linearly interpolate between Strike_{i,j} and
         Strike_AIP at time Settlement_Term to obtain the
         Strike_at_Expiry_j
      Linearly interpolate between B_{i,j} and B_{i+1,j} to obtain
         B_at_Expiry_j
   Next state j
Else (here option expiry lies prior to time i)
   For state j = −(i−1) to i−1 Step 2
      Strike_{i−1,j} = 1/2 d_{i,j} (Strike_{i,j+1} + Strike_{i,j−1})
   Next state j
   For state j = −(i−1) to i−1 Step 2
      Linearly interpolate between B_{i−1,j} and B_{i,j} to obtain
         B_at_Expiry_j
      Linearly interpolate between Strike_{i−1,j} and Strike_{i,j} to obtain
         Strike_at_Expiry_j
   Next state j
End If
```

Initialise option value at option expiry.

```
For state j at option expiry
   V_at_Expiry_j = Max{B_at_Expiry_j − Strike_at_Expiry_j , 0}
Next state j
```

Determine option value at node directly prior to option expiry.
Let k be the node directly prior to option expiry. This corresponds to either node i or node $i-1$ above.
For state $j = -k$ to k Step 2
  $V_{k,j} = \frac{1}{2}\, d_{k,j} \big(\text{V\_at\_Expiry}_{j+1} + \text{V\_at\_Expiry}_{j-1}\big)$
Next state j

Using backward induction, determine initial option value.
For time $i = k-1$ to 0 Step -1
  For state $j = -i$ to i Step 2
    $V_{i,j} = \frac{1}{2}\, d_{i,j} \big(V_{i+1,j+1} + V_{i+1,j-1}\big)$
  Next state j
Next time i

CHAPTER 15

Calibration of the Heath, Jarrow and Morton framework

The HJM framework demonstrates that when the money market account is used as numeraire, forward rates are not martingales but have a non-zero drift term fully determined by the forward rate volatility. Hence, only the forward rate volatilities are needed to price and hedge interest rate contingent claims. To begin an implementation of HJM, the form of the forward rate volatility function must be specified. Consider the differential form of equation (11.32), the forward rate process under the martingale measure[1]:

$$df(t,T) = \sigma(t,T)\int_t^T \sigma(t,y)dy + \sigma(t,T)d\tilde{z}(t) \tag{15.1}$$

Here the drift at time t of the T maturity instantaneous rate depends on the volatilities of the time t forward rates for all maturities from t to T. To price a claim contingent on forward rates over some interval, say $\{f(t,x),\ x \in [T_1, T_2]\}$, we require knowledge of the time t forward rate volatilities for all maturities t to T_2. To fully value the claim, these forward rates will need to be evolved from time t to expiry of the claim at time T, $T \leq T_1, T_2$. Hence a full continuum of forward rate volatilities $\{\sigma(x,y),\ x \in [t,T],\ y \in [t,T_2]\}$ is required. Such demanding data requirements quickly make the valuation impractical or intractable, hence constraints need to be imposed on the evolution of the forward rate volatilities.

This may be done by imposing a functional form on the volatility functions, where each unique functional form leads to a unique HJM model. The volatility functions may be estimated using two distinct approaches:

- Historical volatility estimation – historical time series of forward rates are used to determine the volatilities,
- Implied volatility estimation – current market prices of vanilla interest rate derivatives (caps and swaptions) are used to obtain volatility functions such that these market prices are as best possible reproduced by the model.

[1] To make the notation less cumbersome, I assume a single source of uncertainty, hence $n = 1$, and $\sigma_1(t,T) \equiv \sigma(t,T)$.

First, we consider possible forms of the volatility function within three broad categories. These groupings are not distinct, but rather depict various characteristics that influence choice of volatility function.

(1) Well known volatility functions.
(2) Gaussian volatility functions, i.e. those giving rise to Gaussian forward rate dynamics. The advantage of such forward rate dynamics is that there exist analytic formulae for prices of certain vanilla options.
(3) Functional forms producing Markovian short rate dynamics. This means that $\sigma(\omega, t, T)$ depends only on the state space (ω) observable at time t. This property allows pricing by means of a recombining lattice where the number of nodes grows linearly with time. This contrasts to non-Markovian models, which need to be implemented by simulation or non-recombining lattices (bushy trees) where the number of nodes increases exponentially with time.

15.1. Volatility Function Specifications

15.1.1. Well known volatility functions. Some simple functional forms, many relating to volatility specifications in earlier short rate models, have been investigated for goodness of fit.

(1) $\sigma(\omega, t, T) \equiv \sigma$ where σ is a constant. This is the volatility function of the Ho-Lee model as examined in Chapter 11 §11.7. It leads to a highly tractable but unrealistic model. In Chapter 11 equation (11.38), the forward rate process is shown to be:

$$f(t, T) = f(0, T) + \sigma^2 t\big(T - \tfrac{1}{2}\, t\big) + \sigma \tilde{z}(t)$$

The constant volatility across maturities is at odds with the declining volatility with increasing maturity observed in most economies. Additionally the drift increases with maturity, causing long rates to become unbounded.

(2) $\sigma(\omega, t, T) \equiv \sigma e^{-\lambda(T-t)}$, where σ and λ are constants. As shown in §11.9.1, this is the forward rate volatility function corresponding to the Vasicek model. The associated short rate volatility is constant, $\sigma(\omega, t, t) \equiv \sigma$, but a mean-reverting drift allows the forward rate volatility to decline exponentially with maturity. This specification leads to a tractable but still somewhat unrealistic model.

(3) $\sigma(\omega, t, T) \equiv \frac{\sigma\sqrt{r_t}\delta^2 e^{\delta(T-t)}}{[\phi(e^{\delta(T-t)}-1)+2\delta]^2}$, where δ and ϕ are constants. As derived in §11.8, this is the forward rate volatility resulting from the CIR model specifications. As demonstrated in Chapter 11 §11.9.2 the short rate is Markovian.

(4) $\sigma(\omega, t, T) \equiv \sigma f,\ \sigma\sqrt{f},\ \sigma_0 + \sigma_1(T-t),\ (\sigma_0 + \sigma_1(T-t))f$. Amin and Morton [**2**] investigated these functional forms as well as those in (1)

and (2) above as special cases of the general volatility function:

$$\sigma(\omega, t, T) = (\sigma_0 + \sigma_1(T-t))\, e^{(-\lambda(T-t))} f(t,T)^\gamma$$

They examined parameter stability, goodness of fit to market prices and possibilities of exploiting any mispricing indicated by the models. Although the two-parameter volatility functions displayed better fit to market prices, the associated parameters were less stable over time and were not as consistent in allowing abnormal profits to be earned from indicated mispricing.

15.1.2. Gaussian volatility functions. The forward rate process (equation (11.32) of Chapter 11) is Gaussian if the volatilities are deterministic functions of t and T only, hence:

$$df(t,T) = \alpha(t,T)dt + \sum_{i=1}^{n} \sigma_i(t,T)\, d\tilde{z}_i(t) \tag{15.2}$$

with $\sigma_i(t,T)$, $i = 1, \ldots, n$ deterministic functions of t and T. This implies $f(t,T)$ is normally distributed and hence discount bond prices are lognormally distributed. This attribute allows certain simple contingent claims to be priced using analytical formulae. The time t price of a T_1 maturity European call option on a discount bond with maturity T_2 and strike X, has solution:

$$\begin{aligned} c(t, T_1, T_2) &= P(t,T_2)N(h) - XP(t,T_1)N(h-\varphi) \\ h &= \frac{1}{\varphi}\ln\left(\frac{P(t,T_2)}{XP(t,T_1)}\right) + \frac{1}{2}\varphi \\ \varphi^2 &= \sum_{i=1}^{n}\int_t^{T_1} \left(a_i(u,T_2) - a_i(u,T_1)\right)^2 du \end{aligned} \tag{15.3}$$

where, as defined in (11.11)

$$a_i(t,T) = -\int_t^T \sigma_i(t,y)dy \qquad i = 1, \ldots, n$$

As with other Gaussian interest rate models, there is a positive probability of negative interest rates, which leaves this specification open to criticism.

15.1.3. Markovian Functional Forms. The conditions that must be satisfied to ensure Markovian short rate dynamics are detailed in §11.9. The specific volatility functions of the Ho-Lee, Vasicek and CIR models satisfy these conditions. Ritchken and Sankarasubramanian (RS) [**48**] and Li, Ritchken and Sankarasubramanian [**37**] recognise that these conditions are restrictive since they may impose undesirable characteristics on the term structure evolution and restrict the shape of initial term structure that may be fitted. They identify a broad class of volatility functions that allow contingent claims to be priced by a two-state variable Markovian model. The path dependence is not fully removed, but captured within an additional variable. This class of volatility functions is of the form:

$$\sigma(\omega, t, T) = \sigma(t, t)\, e^{-\int_t^T \kappa(s)ds} \tag{15.4}$$

where $\sigma(t,t)$ is the short rate volatility and $\kappa(t)$ is an exogenously specified deterministic function. RS allow a second variable:

$$\phi(t) = \int_0^t \sigma^2(u,t)du \tag{15.5}$$

which represents the accumulated variance up to date t and houses the path dependence. The link between forward rate volatilities and the short rate volatility is explicitly defined by the parameter $\{\kappa(x)|x \geq 0\}$ and so forward rate volatilities may fluctuate with the level of the short rate. However, forward rate dependent volatilities (as investigated by Amin and Morton [**2**]) are excluded from this framework.

No restrictions are placed on the short rate volatility which could depend on the entire history of term structures, hence the short rate volatility may be a function of both $r(t)$ and $\phi(t)$. Specifically consider

$$\sigma(t,t) = \sigma\,(r(t))^{\gamma}$$

Allowing $\gamma = 0$ and $\gamma = 0.5$ results in special cases of this class corresponding to the Vasicek [**50**] and CIR [**18**] models.

The Hull–White extended Vasicek model (Chapter 7) displays short rate dynamics and forward rate volatility of the form:

$$\begin{aligned} dr(t) &= (\theta(t) - a(t)r(t))\,dt + \sigma(t)\,dz(t) \\ \sigma(\omega, t, T) &= \sigma(t)\,e^{-\int_t^T a(s)ds} \end{aligned}$$

which clearly fits within the RS framework.

Forward rate volatility term structures implied by market prices of traded instruments often display a humped shape, being higher for shorter maturities, see Heath [**26**] and Brigo and Mercurio [**12**]. This volatility structure may be achieved within the RS framework by choosing $\kappa(x)$ to be of an appropriate form. Alternatively, Mercurio and Moraleda (see [**12**]) model a humped volatility structure by specifying the forward rate volatility as:

$$\sigma(\omega, t, T) = \sigma(1 + \gamma(T - t))\, e^{-\lambda(T-t)} \tag{15.6}$$

However, this volatility is not of RS form.

15.2. Implied Volatility Specification

Market prices of traded interest rate derivatives are used to estimate the volatility functions. The choice of number and form of the volatility functions is driven by the users' discretion, desire for the model to exhibit certain characteristics and quality of data available. It is common practice to assume a functional form for the volatilities and then determine the parameters such that the market prices are matched as closely as possible. Consider the volatility functions tested by Amin and Morton [**2**]. These are special cases of the general functional form:

$$\sigma(\omega, t, T) = (\sigma_0 + \sigma_1(T - t))\, e^{(-\lambda(T-t))} f(t, T)^{\gamma}$$

which has four parameters σ_0, σ_1, λ and γ. Amin and Morton found that attempting to calibrate a general functional form, i.e. one with more than two parameters, resulted in unstable parameter values with large estimation errors. They found that given the data available, calibration with two parameters proved optimal. Once a suitable parameterisation is chosen, the parameter values are estimated such that model prices best match market prices. This may be done by a simple procedure such as minimising the sum of squared errors, that is:

$$\begin{aligned} \text{Solve for} \quad & \Theta = (\theta_1, \ldots, \theta_m) \\ \text{such that} \quad & \min \sum_{i=1}^{k} \left[V_i(\Theta) - \bar{V}_i \right]^2 \end{aligned}$$

where

Θ	–	vector of parameters specifying the volatility function,
$V_i(\Theta)$	–	model price of interest rate derivative i, based on parameter values Θ,
$\bar{V}_i$	–	market price of interest rate derivative i,
k	–	number of interest rate derivatives used in calibration on a given day.

15.3. Historical Volatility Specification

Calibration to historical data could be performed in a manner similar to that described for implied volatility data. First a specific formulation of the HJM model is chosen, that is a specific volatility structure is imposed by specifying the number of volatility factors and their specific functional form. Time series of historical term structure data is used to determine the function parameters such that the specified functions fit the data as closely as possible. Such a restrictive approach may be desirable since a specific structure is imposed on the data, allowing for analytical prices of vanilla options. However, such a calibration procedure will not produce an exact match to market prices.

An approach more often applied to historical calibration is that of principal components which allows the number and specific structure of the volatility factors to be directly implied from the data.

15.4. Principal Component Analysis

Principal component analysis (PCA) may be performed on a time series of historical term structure data in an attempt to identify the dominant factors driving its evolution. PCA produces factors maximising successive contributions to overall variance. Hence these factors attempt to explain the diagonal of the covariance matrix. The resulting factors are surrogate volatility factors derived from an empirical analysis of term structure data.

Principal component analysis provides a direct indication of the number and general shape of factors driving the term structure movements. A historical estimate of the magnitude of the volatility functions is also obtained as part of the analysis. These driving factors are both econometrically and financially justifiable, but like all historical calibration methodologies, will not exactly recover market prices of traded derivative instruments.

Many analyses have used spot interest rates as a description of the term structure; here we consider a finite set of forward rates of predetermined tenor, that span the whole term structure. Within the HJM framework each instantaneous forward rate is a stochastic variable in its own right, displaying some degree of correlation with other forward rates. To model a realistic evolution of the term structure one needs to determine a set of forward rate variances (volatilities) as well as covariances (correlations) between the forward rates.

15.4.1. Methodology of principal component analysis. On each day t we observe a vector $f(t)$, of N forward rates with maturities $t+j,\ j = 1,\dots,N$, that is $f(t) = [f(t,t+1)\ f(t,t+2)\ \dots\ f(t,t+N)]'$. A discrete time representation of the forward rate evolution depicted in equation (15.2) is:

$$f(t+\Delta,t+j) \approx f(t,t+j) + \alpha(t,t+j)\Delta + \sum_{i=1}^{N} \sigma_i(t,t+j)\Delta z_i(t) \tag{15.7}$$

Since we have N different forward rates, we may have at most N sources of uncertainty.

15.4.1.1. *Volatility Component.* Choice of two volatility functions is possible:

- Gaussian volatility function, where

  $$\sigma_i(t,T) = \sigma_i(T-t)$$

  is a deterministic function of term to maturity $T-t$, $\forall i = 1,\dots,N$.
- Lognormal volatility function, that is

  $$\sigma_i(t,T) = \sigma_i(T-t)\min\{f(t,T), M\}$$

  where $\sigma_i(T-t)$ is a deterministic function of term to maturity $T-t$, $\forall i = 1,\dots,N$, and M is some large positive constant[2].

15.4.1.2. *Drift component.* The drift vector in (15.7), $\alpha = [\alpha(t,t+1)\ \alpha(t,t+2)\ \dots\ \alpha(t,t+N)]$ can be determined from the historical data. For purposes of pricing contingent claims, we derive the drift from the no arbitrage condition such that it is consistent with the risk-neutral valuation framework. This implies that for purposes of determining the covariance matrix, we may assume $\alpha = 0$. Alternatively, mean adjusted changes in forward rates may be used, thereby isolating the unexpected changes in forward rates[3]. Since PCA deals with unexpected changes in the term structure, the drift does not enter the calculations and the components obtained from the two methodologies are consistent.

15.4.1.3. *Calculation Methodology.* From (15.7), we define the $(N \times 1)$ vector $d(t) = [d_1(t)\ d_2(t)\ \dots\ d_N(t)]'$ where each element is constructed as:

$$\begin{aligned} d_j(t) &= \frac{f(t+\Delta t, t+j) - f(t,t+j)}{\sqrt{\Delta t}} - \alpha(t,t+j)\sqrt{\Delta t} \\ &= \sum_{i=1}^{N} \sigma_i(t,t+j)\frac{\Delta z_i(t)}{\sqrt{\Delta t}}, \quad j = 1,\dots,N \end{aligned} \tag{15.8}$$

where $\frac{\Delta z_i(t)}{\sqrt{\Delta t}}$, $i = 1,\dots,N$ are standard normal. $d_j(t)$ is the mean-adjusted change in forward rate with maturity $t+j$ over time period $[t, t+\Delta t]$ and represents a realisation of the random variable d_j, the change in the j maturity forward rate over some time period Δt.

For a set of historical time-series data with K observations of vector $d(t_k)$, $k = 1,\dots,K$ we define[4]:

[2] This is consistent with the discussion in Chapter 11 §11.5 concerning the lognormality of forward rates.

[3] This is the methodology used here

[4] This simplified calculation of the covariance matrix implicitly assumes matrix D is mean adjusted and cannot be applied if unadjusted changes in forward rates are used.

$$\Sigma = \frac{1}{K-1} DD' = \text{covar}(D) \tag{15.9}$$

where D is a $(N \times K)$ matrix with columns $d(t_k),\ k = 1, \ldots, K$ and Σ is the $(N \times N)$ covariance matrix with elements[5] $\Sigma_{i,j} = \text{covar}(d_i, d_j)$. Since this covariance matrix is symmetrical and positive semi-definite, we can find an orthogonal[6] matrix A such that:

$$\begin{aligned} \Sigma &= A \Lambda A' \\ \Rightarrow \Lambda &= A' \Sigma A \end{aligned} \tag{15.10}$$

where Λ is an $(N \times N)$ diagonal matrix with elements $\lambda_m,\ m = 1, \ldots, N$ such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. The columns of the $(N \times N)$ matrix $A = [a_1,\ a_2,\ \ldots,\ a_N]$ are the N eigenvectors of covariance matrix Σ, while the N diagonal elements matrix Λ are the corresponding eigenvalues.

15.4.2. Interpretation of the Principal Components. Let us examine what we have achieved by the above transformation. The starting point is an $(N \times K)$ matrix $D = [d(t_1),\ d(t_2),\ \ldots,\ d(t_K)]$, its columns being K observations the original vector of variables $d(\cdot) = [d_1(\cdot)\ d_2(\cdot)\ \ldots\ d_N(\cdot)]'$. Multiplication by the transpose of the orthogonal matrix A transforms D into a new matrix G:

$$G = A'D \tag{15.11}$$

where $G = [g(t_1),\ g(t_2),\ \ldots,\ g(t_K)]$ is an $(N \times K)$ matrix with columns being realisations of the vector of new variables $g(\cdot) = [g_1(\cdot)\ g_2(\cdot)\ \ldots\ g_N(\cdot)]'$. We have defined the new variables as the weighted sums of the original variables, the weights being the elements of the corresponding eigenvectors. Hence the new variable g_i is the weighted sum of all original variables $d_j,\ j = 1, \ldots, N$, weighted by the elements of the i^{th} eigenvector a_i:

$$g_i(t_k) = \sum_{j=1}^{N} a_{ji}\, d_j(t_k) \qquad \text{for each observation time } k,\ k = 1, \ldots, K.$$

[5] Here i and j represent forward rate maturities.

[6] An orthogonal or Hermitian matrix is such that its transpose is equal to its inverse i.e. $A' = A^{-1}$.

Making use of (15.9), (15.10) and (15.11) the definition of G, we find:

$$\begin{aligned}
\operatorname{covar}(G) &= \operatorname{covar}(A'D) \\
&= A' \operatorname{covar}(D)\, A \\
&= A'\,\Sigma\, A \\
&= A'\, A\Lambda A'\, A \\
&= \Lambda
\end{aligned}$$

Therefore the diagonal matrix of eigenvalues Λ, is the covariance matrix of the new variables. Since the transformed variables have zero covariance, they are orthogonal (statistically independent). Originally, the stochastic evolution of each forward rate is driven by a separate source of uncertainty which has some degree of correlation with all other sources of uncertainty driving other points on the yield curve. We have transformed to a set of factors, each affecting every point on the yield curve, the extent to which each forward rate is affected is determined by the eigenvector. By virtue of matrix A's orthogonality we have:

$$G = A'D \qquad \Rightarrow D = AG$$

and so each original variable $d_j,\ j = 1, \ldots, N$ may be written as:

$$d_j(t_k) = \sum_{i=1}^{N} a_{ji}\, g_i(t_k) \qquad \text{for each observation time } k,\ k = 1, \ldots, K$$

A move in the i^{th} new variable g_i influences each of the original variables $d_j,\ j = 1, \ldots, N$ by a factor of a_{ji}.

Let $\Lambda^{\frac{1}{2}} = \operatorname{diag}[\sqrt{\lambda_1},\ \sqrt{\lambda_2},\ \ldots,\ \sqrt{\lambda_N}]$ and define a matrix $W = [w(t_1),\ w(t_2),\ \ldots,\ w(t_K)]$ such that:

$$\begin{aligned}
W &= \Lambda^{-\frac{1}{2}} G \\
\Rightarrow \operatorname{covar}(W) &= \operatorname{covar}(\Lambda^{-\frac{1}{2}} G) \\
&= \Lambda^{-\frac{1}{2}} \operatorname{covar}(G) \Lambda^{-\frac{1}{2}} \\
&= \Lambda^{-\frac{1}{2}}\, \Lambda\, \Lambda^{-\frac{1}{2}} \\
&= I
\end{aligned}$$

So W is a matrix of independent standard Brownian variables. We may now express the matrix of original variables D, in terms of this matrix of standard Brownian variables as:

$$\begin{aligned}
D &= AG \\
&= A\, \Lambda^{\frac{1}{2}}\, W
\end{aligned} \tag{15.12}$$

or algebraically we may write each $d_j, \ j = 1, \ldots, N$ as:

$$d_j(t_k) = \sum_{i=1}^{N} a_{ji} \sqrt{\lambda_i}\, w_i(t_k) \qquad \text{for each observation time } k, \ k = 1, \ldots, K$$

which by the definition of $d_j(t_k)$ in (15.8), allows us to write the forward rate process as:

$$\begin{aligned} f(t + \Delta t, t + j) - f(t, t + j) &= \alpha(t, t + j)\Delta t + \sum_{i=1}^{N} a_i(t + j) \sqrt{\lambda_i} \sqrt{\Delta t}\, w_i(t) \\ &= \alpha(t, t + j)\Delta t + \sum_{i=1}^{N} a_i(t + j) \sqrt{\lambda_i}\, \Delta w_i(t) \end{aligned}$$

or in continuous time

$$df(t, t + j) = \alpha(t, t + j)dt + \sum_{i=1}^{N} a_i(t + j) \sqrt{\lambda_i}\, dw_i(t) \tag{15.13}$$

where $a_i(t + j) = a_{ji}$ for $j = 1, \ldots, N$.

From equation (15.12) we have:

$$\text{cov}(D) = \text{cov}(A\Lambda^{\frac{1}{2}} W) = A\Lambda^{\frac{1}{2}} (A\Lambda^{\frac{1}{2}})'$$

and $A\Lambda^{\frac{1}{2}}$ is the matrix of principal components of covariance matrix Σ.

As mentioned before, the principal components maximise successive contributions to overall variance. Also, we know $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$, with the first principal component corresponding to the largest eigenvalue. Therefore the first principal component accounts for the largest percentage of total variance, the second component accounts for the largest percentage of the remaining (residual) variance and so on. The greater the covariance between points on the yield curve, the larger the percentage of total variance that will be explained by the first few principal components. So in general, the more principal components required to explain the total variance, the lower the covariance across the yield curve.

While the number of principal components obtained from the analysis detailed in the previous section equals the number of different maturity forward rates used, the hope is that just a few dominant principal components will be required to explain the majority of the variance.

15.5. Choosing the number of volatility factors

The HJM framework has the advantage that it allows seamless extension from one to several sources of uncertainty. The choice of number of volatility factors to use is driven by a number of often conflicting considerations. Additional sources of uncertainty introduce additional degrees of freedom to the evolution of the term structure which allows decorrelation (decreasing correlation between forward rates with larger differences in their maturities) of rates to be properly accounted for. However, additional factors increase the numerical complexity, slowing down computation time.

Even when calibrating via PCA, where this choice is driven by the number of factors explaining a sufficiently large percentage of total volatility, there are several points to consider:

(1) PCA tells us how many factors are significant in explaining the movement of the term structure within the historical data set. We wish to use these factors as predictors of future movements of the term structure. While adding additional factors will improve our ability to explain the historical term structure movements, this may not be the case for explaining future movements.
(2) The historical factors and their specific loadings (magnitude for various maturities along the term structure) may not be ideal predictors of future volatility factors [**52**]. Additional factors may not be stable through time adding noise, rather than improving the structural relationships within the volatility predictions.
(3) PCA attempts to maximise the factors' contribution to the diagonal elements of the covariance matrix, that is the variances. The resulting factors may explain the off-diagonal elements (covariances) with significant error. This may affect valuation of correlation dependent options such as swaptions and spread options.

15.6. Combining Historical and Implicit Calibration

While the volatility level changes frequently, correlations between various maturity forward rates tend to be more stable. This lends itself to extracting the correlation term structure from historical data and matching the exact level of volatility to market implied volatilities. The volatility associated with factor i, $\sigma_i(t, t+j)$, may be expressed as:

$$\sigma_i(t, t+j) = \sigma(t, t+j)\psi_i(t, t+j)$$

where $\sigma(t, t+j)$ is the total volatility of forward rate with maturity $t+j$ and $\psi_i(t, t+j)$ is the weight associated with factor i. These functions $\psi_i(t, \cdot)$, are elements of the i^{th} principal component of the correlation matrix; hence[7]:

$$\sum_{i=1}^{N} \psi_i(t, t+j)^2 = 1$$

$$\sum_{i=1}^{N} \psi_i(t, t+j)\psi_i(t, t+k)$$

are the diagonal and off diagonal elements of the correlation matrix respectively. The forward rate volatilities $\sigma(t, t+j)$, are the market implied Black volatilities for caplets expiring at time $t+j$.

[7] As in (15.10), we decompose the correlation matrix C such that $C = A\Lambda A'$ and hence define the matrix of principle components as $\Psi = A\Lambda^{\frac{1}{2}}$ where $\Psi\Psi' = C$

Closing Remarks

Initial attempts to model the term structure of interest rates developed along the same lines as stock price models. These term structure models are based within an economic setting where Brownian motions give rise to random shocks. Most models make the assumption that the entire term structure is driven by the short-term interest rate of interest. Some models allow this short-term interest rate to be driven by underlying economic variables, hence introducing multiple factors. Other models introduce a second factor such a long interest rate or short-term interest rate volatility, modelling these two factors as mutually dependent processes.

The early models are concerned with determining an appropriate level of the term structure in such a way that it is consistent with the underlying economic model. This makes it difficult to incorporate information from an initial observed term structure and hence to reproduce the market prices of securities.

A change of perspective introduced the need for a model to perfectly fit an initial term structure and reproduce market-observed prices of vanilla instruments. The focus shifted to calibrating model parameters in such a way as to account for, rather than explain, the shape of the yield curve. This has remained the driving force behind current research into term structure modelling. Given the market-observed prices of vanilla securities, practitioners need to price more exotic instruments in a consistent manner.

Models allowing the instantaneous short-term interest rate to be the single driving factor of the entire yield curve are quite restrictive, in part because returns on bonds of all maturities are instantaneously perfectly correlated. This is unrealistic and imposes restrictions on the resulting yield curve which makes many market-observed term structures difficult, if not impossible, to replicate.

The HJM framework introduces a new perspective to term structure modelling. By allowing the instantaneous forward rate to be the fundamental variable, they are able to specify the entire term structure at any one time. This is in contrast to models where the instantaneous short-term interest rate, a single point on the yield curve, is the fundamental variable.

However, the HJM approach still shares a fundamental problem with all its predecessors. The state variables are in fact unobservable: instantaneous short and/or forward interest rates do not trade in the market. Hence to calibrate these models one must perform a translation of unobservable model variables to appropriately selected market proxies. Among others, this is one of the landmark features of the approach taken by BGM, who develop a model which determines the stochastic process followed by a market traded rate of interest: the discretely compounded LIBOR. This introduces yet another dimension to term structure models since a trader may directly express her/his views on movements in market traded quantities. The BGM model becomes a tool whereby a trader's views are directly translated into option prices.

The older more traditional models such as Vasicek and CIR still have a place in the financial markets. Movements in market variables, that cannot be replicated within the model, will show up as anomalies and mispricing. This may lead users to perform a more detailed analysis of the causes of such anomalies. Therefore, the qualitative insight they provide about the dynamics of the yield curve can be beneficial for the understanding of more advanced models.

Bibliography

[1] K. Adams, *Smooth interpolation of zero curves*, Algo Research Quarterly **4** (2001), no. 1/2, 11–22.

[2] K. I. Amin and A. J. Morton, *Implied volatility functions in arbitrage-free term structure models*, Journal of Financial Economics **35** (1994), 141–180.

[3] L. Arnold, *Stochastic differential equations*, John Wiley and Sons Ltd., 1974.

[4] C. A. Ball and W. N. Torous, *Bond price dynamics and options*, Journal of Financial and Quantitative Analysis **18** (1983), 517–531.

[5] F. Black, *The pricing of commodity contracts*, Journal of Financial Economics **3** (1976), 167–179.

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

[7] F. Black and P. Karasinski, *Bond and option pricing when short rates are lognormal*, Financial Analysts Journal (1991), 52–59.

[8] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), 637–659.

[9] A. Brace, D. Gaterek, and M. Musiela, *The market model of interest rate dynamics*, Mathematical Finance **7** (1997), 127–154.

[10] M. J. Brennan and E. S. Schwartz, *A continuous time approach to the pricing of bonds*, Journal of Banking and Finance **31** (1979), 33–155.

[11] ———, *An equilibrium model of bond pricing and a test of market efficiency*, Journal of Financial and Quantitative Analysis **17** (1982), 301–329.

[12] D. Brigo and F. Mercurio, *Interest rate models theory and practice*, Springer-Verlag, 2001.

[13] Richard L. Burden and J. Douglas Faires, *Numerical analysis*, sixth ed., Brooks Cole, 1997.

[14] A. Carverhill, *When is the short rate Markovian?*, Mathematical Finance **4** (1994), 305–312.

[15] Les Clewlow and Chris Strickland, *Implementing derivatives models*, John Wiley and Sons Ltd., 1998.

[16] Donald L. Cohn, *Measure theory*, Birkhauser, 1980.

[17] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *An intertemporal general equilibrium model of asset prices*, Econometrica **53** (1985), 363–384.

[18] ———, *A theory of the term structure of interest rates*, Econometrica **53** (1985), 385–407.

[19] J. C. Cox, S. A. Ross, and M. Rubinstein, *Option pricing: A simplified approach*, Journal of Financial Economics **7** (1979).

[20] Carl de Boor, *A practical guide to splines*, Springer-Verlag New York Inc., 1978.

[21] R.J. Elliott, *Stochastic calculus and applications*, Springer-Verlag, 1982.

[22] H. Geman, El Karoui N., and J. Rochet, *Changes of numeraire, changes of probability measure and option pricing*, Journal of Applied Probability **32** (1995).

[23] J. M. Harrison and D. M. Kreps, *Martingales and arbitrage in multiperiod securities markets*, Journal of Economic Theory **20** (1979), 381–408.

[24] J. M. Harrison and S. R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Processes and their Application **11** (1981), 215–260.

[25] D. Heath, R. Jarrow, and A. Morton, *Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation*, Econometrica **60** (1992).

[26] D. Heath, R. Jarrow, A. Morton, and M. Spindel, *Easier done than said*, Risk Magazine **5** (1992), no. 9.

[27] T. S. Y. Ho and S.-B. Lee, *Term structure movements and pricing interest rate contingent claims*, Journal of Finance **41** (1986).

[28] J. Hull and A. White, *Pricing interest rate derivative securities*, Review of Financial Studies **3** (1990).

[29] J. C. Hull and A. White, *Hull-White on derivatives - a compilation of articles*, Risk Publications, 1996.

[30] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland/Kodansha, 1981.

[31] F. Jamshidian, *An exact bond option formula*, Journal of Finance **44** (1989).

[32] A. Jeffrey, *Single factor Heath-Jarrow-Morton term structure models based on the Markov spot interest rate dynamics*, Journal of Financial and Quantitative Analysis **30** (1995).

[33] I. Karatzas and S.E. Shreve, *Brownian motion and stochastic calculus*, Springer-Verlag, 1988.

[34] S. Karlin and H.M. Taylor, *A first course in stochastic processes*, Academic Press, 1975.

[35] H. O. Lancaster, *The chi-squared distribution*, John Wiley and Sons Ltd., 1969.

[36] T. C. Langetieg, *A multivariate model of the term structure*, Journal of Finance **35** (1980).

[37] A Li, P. Ritchken, and L. Sankarasubramanian, *Lattice models for pricing American interest rate claims*, Journal of Finance **50** (1995).

[38] F.A Longstaff and E.S. Schwartz, *Interest rate volatility and the term structure: A two-factor general equilibrium model*, Journal of Finance **47** (1992).

[39] ———, *A two-factor interest rate model and contingent claims valuation*, Journal of Fixed Income (1992).

[40] Robert C. Merton, *Optimum consumption and portfolio rules in a continuou-time model*, Journal of Economic Theory **3** (1971).

[41] M. Musiela and M. Rutkowski, *Martingale methods in financial modelling*, Springer-Verlag, 1997.

[42] R.W.D. Nickalls, *A new approach to solving the cubic: Cardan's solution revealed*, The Mathematical Gazette **77** (1993).

[43] B. Oksendal, *Stochastic differential equations*, fifth ed., Springer-Verlag, 1998.

[44] M. Parkin and D. King, *Economics*, Addison-Wesley, 1995.

[45] R. Rebonato, *Interest-rate option models*, John Wiley and Sons Ltd., 1996.

[46] ———, *Calibrating the BGM model*, Risk **12** (1999), no. 3.

[47] W. T. Reid, *Riccati differential equations*, Academic Press, 1972.

[48] P. Ritchken and L. Sankarasubramanian, *Volatility structures of forward rates and the dynamics of the term structure*, Mathematical Finance **5** (1995), no. 1.

[49] S. E. Shreve, *Stochastic calculus and finance, lecture notes*, 1997.

[50] O. A. Vasicek, *An equilibrium characterisation of the term structure*, Journal of Financial Economics **15** (1977).

[51] P. Wilmott, *Derivatives: The theory and practice of financial engineering*, John Wiley and Sons Ltd., 1998.

[52] T. Wilson, *Debunking the myths*, Risk **7** (1994), no. 4.

Index