Interest-Rate Models

Andrew J.G. Cairns
Actuarial Mathematics and Statistics
School of Mathematical and Computer Sciences
Heriot-Watt University
Edinburgh, EH14 4AS, United Kingdom

E-mail: A.Cairns@ma.hw.ac.uk
WWW: http://www.ma.hw.ac.uk/∼andrewc/

Prepared for the Encyclopaedia of Actuarial Science

1 Introduction

In this article we will describe some of the main developments in interest-rate modelling since Black & Scholes' (1973) and Merton's (1973) original articles on the pricing of equity derivatives. In particular, we will focus on continuous-time, arbitrage-free models for the full *term structure of interest rates*. Other models which model a limited number of key interest rates or which operate in discrete time (for example, the Wilkie (1995) model) will be considered elsewhere. Additionally, more detailed accounts of *affine term-structure models* and *market models* are given elsewhere in this volume.

Here we will describe the basic principles of arbitrage-free pricing and cover various frameworks for modelling: short-rate models (for example, Vasicek, Cox-Ingersoll-Ross, Hull-White); the Heath-Jarrow-Morton approach for modelling the forward-rate curve; pricing using state-price deflators including the Flesaker-Hughston/Potential approach; and the Markov-functional approach.

The article works through various approaches and models in a historical sequence. Partly this is for history's sake, but, more importantly, the older models are simpler and easier to understand. This will allow us to build up gradually to the more up to date, but more complex, modelling techniques.

1.1 Interest rates and prices

One of the first problems one encounters in this field is the variety of different ways of presenting information about the term structure.[1] The expression "yield curve" is often used in a sloppy way with the result that it often means different things to different people: how is the yield defined; is the rate annualised or semi-annual or continuously compounding; does it refer to the yield on a coupon bond or a zero-coupon bond? To avoid further confusion then, we will give some precise definitions.

- We will consider here only default-free government debt. Bonds which involve a degree of credit risk will be dealt with in a separate article.
- The basic building blocks from the mathematical point of view are zero-coupon bonds.[2] In its standard form such a contract promises to pay £1 on a fixed date in the future. Thus we use the notation $D(t, T)$ to represent the value at time t of £1 at time T.[3] The bond price process has the boundary conditions $D(T, T) = 1$ and $D(t, T) > 0$ for all $t \leq T$.
- A *fixed-income* contract equates to a collection of zero-coupon bonds. For example, suppose it is currently time 0 and the contract promises to pay the fixed amounts $c_1, c_2, \ldots, c_n$ at the fixed times $t_1, t_2, \ldots, t_n$. If we assume that there are no taxes[4] then the fair or market price for this contract at time 0 is

$$P = \sum_{i=1}^{n} c_i D(0, t_i).$$

(This identity follows from a simple, static hedging strategy which involves replicating the coupon bond payments with the payments arising from a portfolio of zero-coupon bonds.) The *gross redemption yield* (or *yield-to-maturity*) is a measure of the average interest rate earned over the term of the contract given the current price P. The gross redemption yield is the

[1] We implicitly assume that readers have gone beyond the assumption that the yield curve is flat!

[2] From the practical point of view it is sensible to start with frequently-traded coupon bonds, the prices of which can be used to back out zero-coupon-bond prices. Zero-coupon bonds do exist in several countries, but they are often relatively illiquid making their quoted prices out of date and unreliable.

[3] Here the $D(\cdot)$ notation uses D for discount bond or discounted price. Common notation used elsewhere is $P(t, T)$ for price and $B(t, T)$ for bond price. Additionally, one or other of the t or T can be found as a subscript to distinguish between the nature of the two variables: t is the dynamic variable, while T is usually static.

[4] Alternatively we can assume that income and capital gains are taxed on the same mark-to-market basis.

solution, δ, to the equation of value

$$P = \hat{P}(\delta) = \sum_{i=1}^{n} c_i e^{-\delta t_i}.$$

If the c_i are all positive then the solution to this equation is unique.

δ as found above is a continuously compunding rate of interest. However, the gross redemption yield is usually quoted as an annual (that is, we quote $i = \exp(\delta) - 1$) or semi-annual rate ($i^{(2)} = 2[\exp(\delta/2) - 1]$) depending on the frequency of contracted payments.[5]

- The *spot-rate curve* at time t refers to the set of gross redemption yields on zero-coupon bonds. The spot rate at time t for a zero-coupon bond maturing at time T is denoted by $R(t,T)$ which is the solution to $D(t,T) = \exp[-R(t,T)(T-t)]$: that is,

$$R(t,T) = \frac{-1}{(T-t)} \log D(t,T).$$

- The instantaneous, *risk-free* rate of interest is the very short-maturity spot rate

$$r(t) = \lim_{T \to t} R(t,T).$$

This gives us the *money-market account* or *cash account* $C(t)$ which invests only at this risk-free rate. Thus $C(t)$ has the stochastic differential equation (SDE)

$$dC(t) = r(t)C(t)dt$$

with solution

$$C(t) = C(0) \exp\left[\int_0^t r(u)du\right].$$

- Spot rates refer to the *on-the-spot* purchase of a zero-coupon bond. In contrast, *forward rates* give us rates of interest which refer to a future period of investment. Standard contracts will refer to both the future delivery and maturity dates. Thus $F(t,T_1,T_2)$ is used to denote the (continuously compounding) rate which will apply between times T_1 and T_2 as determined by a contract entered into at time t. The standard contract also requires that the value of the contract at time t is zero. Thus, a simple no-arbitrage

[5] Thus $\exp(\delta) \equiv 1 + i \equiv \left(1 + \frac{1}{2} i^{(2)}\right)^2$.

argument[6] shows that we must have

$$F(t,T_1,T_2) = \frac{\log\left(D(t,T_1)/D(t,T_2)\right)}{T_2 - T_1}. \tag{1}$$

From the theoretical point of view these forward rates give us too much information. Sufficient information is provided by the *instantaneous forward-rate curve*

$$f(t,T) = \lim_{S\to T} F(t,T,S) = -\frac{\partial}{\partial T}\log D(t,T) = \frac{-1}{D(t,T)}\frac{\partial D(t,T)}{\partial T}.$$

- The *par-yield curve* is a well-defined version of the "yield curve". For simplicity suppose that we have the possibility to issue (without influencing other market prices) a set of coupon bonds, each with its own coupon rate (payable annually) and with maturity at times $t+1, t+2, \ldots$. For each maturity date we ask the question: at what level do we need to set the coupon rate in order that the market price for the coupon bond will be at par? Thus let g_n be the general coupon rate for maturity at $t+n$ for a coupon bond with a *nominal* or *face value* of 100. The market price for this bond will be

  $$P_n(t) = g_n \sum_{i=1}^{n} D(t,t_i) + 100D(t,t+n)$$

  If the bond is currently priced *at par* then this means that $P_n(t) = 100$. This in turn implies that

  $$g_n \equiv \rho(t,t+n) = 100\frac{\big(1 - D(t,t+n)\big)}{\sum_{i=1}^{n} D(t,t+i)}$$

  and $\rho(t,t+n)$ is referred to as the par yield, which we interpret here as an annualised rate.

- Associated with forward rates and par yields we have *LIBOR* and swap rates. Both rates apply to actual rates of interest which are agreed between banks rather than being derived from the government bonds market. LIBOR rates are equivalent to spot rates but are quoted as simple rates of interest: that is, for the term-τ LIBOR an investment of £1 at time t will return $1+\tau L$ at time $t+\tau$. Typically the *tenor* τ is 3, 6 or 12 months. Thus, generalising the notation to $L(t,t+\tau)$ we have

  $$1 + \tau L(t,t+\tau) = \frac{1}{D(t,t+\tau)} = \exp\big(R(t,t+\tau)\big).$$

[6] A short investment of 1 T_1-bond and a long investment of $D(t,T_1)/D(t,T_2)$ T_2-bonds has value zero at time t and produces exactly the same cashflows at T_1 and T_2 as the forward contract with the forward rate defined in equation (1). The law of one price (principle of no arbitrage) tells us that the forward contract must have the same value at time t as the static portfolio of T_1 and T_2 bonds: that is, zero value. Therefore equation (1) gives us the fair forward rate.

> An interest-rate swap contract (with annual payment dates, for simplicity, and a term to maturity of n) is a contract which involves swapping a series of fixed payments for the floating LIBOR rate. The two investors A and B enter into a contract at time t where A pays B the fixed amount $K(t, t+n)$ at times $t+1, t+2, \ldots, t+n$ and in return B pays A $L(t, t+1)$ at time $t+1$ (which is known at time t), $L(t+1, t+2)$ at time $t+2$ (not known until time $t+1$), and so on up to $L(t+n-1, t+n)$ at time t (not known until time $t+n-1$). The fair swap rate $K(t, t+n)$ is the rate at which the value of the contract has zero value at time t and this is very closely linked to par yields.[7]

An example of how some of the different rates of interest interact can be seen in Figure 1 for UK government bonds at the close of business on 31 December 2002. On that date the Bank of England base rate was 4% with declining yields over the first year of maturity before yields climb back up again. Coupon-bond yields do not define a crisp curve: instead two bonds with the same maturity date but different coupon rates can have different gross redemption yields depending upon the shape of the spot-rate curve. We can also see that the spot-rate and par-yield curves are not especially smooth and this reflects a mixture of local supply and demand factors, liquidity and transaction costs. We can calculate 6-month forward rates based on the spot-rates in Figure 1 but this results in a far rougher curve than we would expect. These curves are straightforward to smooth (see, for example, Nelson & Siegel (1987), Svensson (1994), Fisher, Nichka & Zervos (1995), Cairns (1998) and Cairns & Pritchard (2001)) if we are only interested in the term structure on a given date rather than the dynamics or if we require a relatively smooth forward-rate curve as the starting point for the Heath-Jarrow-Morton framework (Section 3).

As an alternative to government bonds data, investment banks often use swap curves as the starting point for pricing many types of derivatives. These tend to be rather smoother because of the high degree of liquidity in the swaps market and because the contracts are highly standardised in contrast to the bonds market which is more heterogeneous. The relative smoothness of the swap curve can be seen in Figure 2. The figure also highlights the relatively wide bid/ask spreads (up to 13 basis points) for interbank swaps in comparison with government bonds (around 1 to 2 basis points) and also the impact of credit risk on swap rates.

We will now look at dynamic, arbitrage-free models for the term-structure of interest rates. For a more detailed development the reader can go to one of several good textbooks which specialise in the subject of interest-rate modelling as well as

[7] Mathematically swap rates and par yields are identical. However, swap rates generally refer to agreements between investment banks whereas par yields derive from the government bonds market. Liquidity and credit considerations then lead to small differences between the two. For example, see Figure 2.

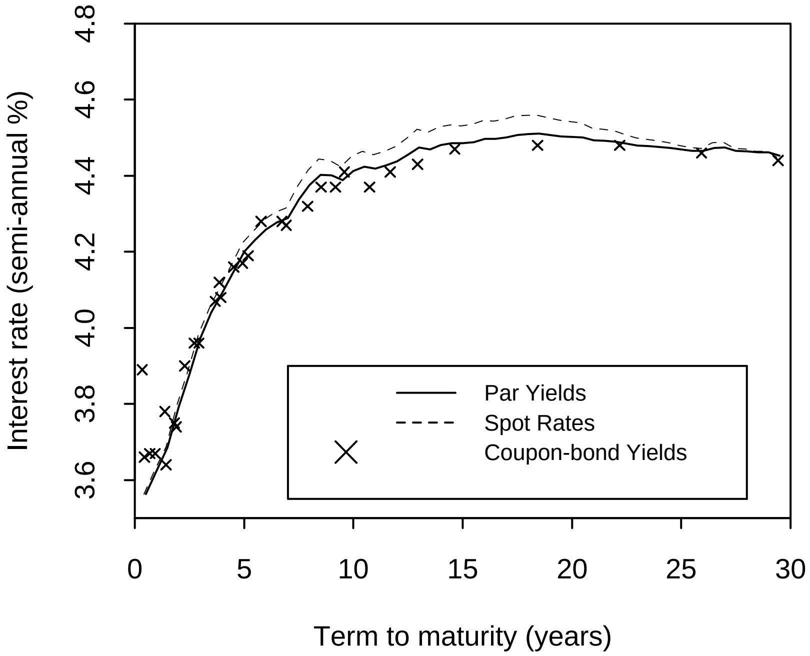

Figure 1: UK Government bond yields at the close of business on 31 December 2002 (source: www.dmo.gov.uk). Coupon bond yields are the gross redemption yields on (single-dated) coupon bonds. The spot-rate curve is derived from STRIPS' prices. The par-yield curve is derived from the spot-rate curve.

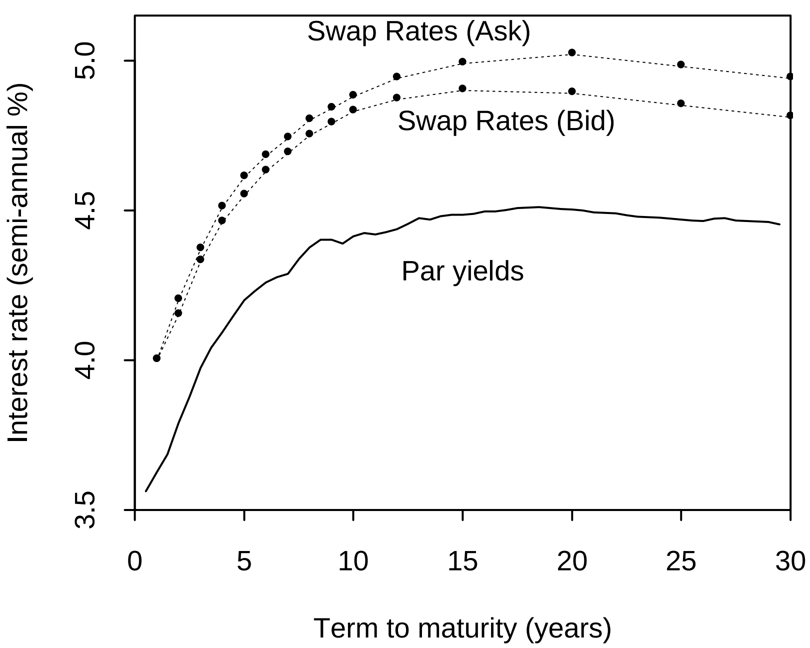

Figure 2: Par yields implied by UK STRIPS market on 31 December 2002 (solid curve) in comparison with interbank swap rates (dotted curves, bid and ask rates) (source Financial Times). Dots represent quoted swap maturity dates.

the original articles referenced here. Those by Brigo & Mercurio (2001), Pelsser (2000), and Rebonato (1998, 2002) are relatively specialised texts which deal principally with the most recent advances. Cairns (2003) and James & Webber (2000) both give a wider overview of the subject: the former starting at a lower level and aimed at persons new to the subject; the latter serving as an excellent and wide-ranging reference book for anyone interested in interest-rate modelling.

2 The risk-neutral approach to pricing

One of the earliest papers to tackle arbitrage-free pricing of bonds and interest-rate derivatives was Vasicek (1997). This paper is best known for the Vasicek model for the risk-free rate of interest, $r(t)$ described below. However, Vasicek also developed a more general approach to pricing which ties in with what we now refer to the as the risk-neutral-pricing approach.

For notational convenience and clarity of exposition, we will restrict ourselves here to one-factor, diffusion models for the short rate, $r(t)$. The approach is, however, easily extended to multifactor models. The general SDE for $r(t)$ is

$$dr(t) = a(t, r(t))dt + b(t, r(t))dW(t)$$

where $W(t)$ is a standard Brownian motion under the *real-world* measure P (also referred to as the *objective* or *physical* measure), and $a(\cdot)$ and $b(\cdot)$ are suitable smooth functions of t and r.

Vasicek also postulated that the zero-coupon prices have the SDE's

$$dD(t,T) = D(t,T)[m(t,T,r(t))dt + S(t,T,r(t))dW(t)]$$

for suitable functions $m(t,T,r(t))$ and $S(t,T,r(t))$ which will depend on the model for $r(t)$. Vasicek first used an argument similar to Black & Scholes (1973) to demonstrate that there must exist a previsible stochastic process $\gamma(t)$ (called the *market price of risk*) such that

$$\frac{m(t,T,r(t)) - r(t)}{S(t,T,r(t))} = \gamma(t)$$

for all $T > t$. Without this strong relationship between the drifts and volatilities of bonds with different terms to maturity the model would admit arbitrage. Second, Vasicek derived a Black-Scholes type of PDE for prices:

$$\frac{\partial D}{\partial t} + (a - b.\gamma)\frac{\partial D}{\partial r} + \frac{1}{2}b^2\frac{\partial^2 D}{\partial r^2} - rD = 0 \quad (2)$$

for $t < T$, with the boundary condition $D(T,T) = 1$. We can then apply the Feynman-Kac formula (see, for example, Cairns (2003)) to derive the well-known risk-neutral pricing formula for bond prices:

$$D(t,T,r(t)) = E_Q\left[\exp\left(-\int_t^T r(u)du\right) \,\middle|\, \mathcal{F}_t\right]. \tag{3}$$

In this formula the expectation is taken with respect to a new probability measure Q rather than the original measure P. Under Q, $r(t)$ has the risk-adjusted SDE

$$dr(t) \quad = \quad \tilde{a}(t,r(t))dt + b(t,r(t))d\tilde{W}(t)$$

where $\tilde{a}(t,r(t)) = a(t,r(t)) - \gamma(t)b(t,r(t))$ and $\tilde{W}(t)$ is a standard Brownian motion under Q.

Additionally, if we let X be some interest-rate derivative payment at some time T then the price for this derivative at some earlier time t is

$$V(t) = E_Q\left[\exp\left(-\int_t^T r(u)du\right) X \,\middle|\, \mathcal{F}_t\right]. \tag{4}$$

The more modern, martingale approach which gives rise to the same result in equation (3) is described in detail in Cairns (2003). Here we first establish the measure Q as that under which the prices of all *tradeable assets*[8] discounted by the cash account, $C(t)$ (that is, the $Z(t,T) = D(t,T)/C(t)$), are martingales. For this reason, Q is also referred to as an *equivalent martingale measure.* In this context $C(t)$ is also described as the *numeraire.* The martingale approach tends to provide a more powerful starting point for calculations and modelling. On the other hand the PDE approach is still useful when it comes to numerical calculation of derivative prices.

In practice, modellers often work in reverse by proposing the model for $r(t)$ first under Q and then by adding the market price of risk to allow us to model interest rates and bond prices under P. Specific models for $r(t)$ under Q include Vasicek (1977)

$$dr(t) = \alpha(\mu - r(t))dt + \sigma d\tilde{W}(t)$$

and Cox, Ingersoll and Ross (1985) (CIR)

$$dr(t) = \alpha(\mu - r(t))dt + \sigma\sqrt{r(t)}d\tilde{W}(t).$$

Both models give rise to analytical formulae for zero-coupon bond prices and European options of the same bonds. The CIR model has the advantage that

[8] Here a *tradeable asset* has a precise meaning in a mathematical sense. It refers to assets where there are no coupon or dividend payments. Thus, the price at time t represents the total return on the investment up to that time with no withdrawals or inputs of cash at intermediate times. For assets which do pay coupons or dividends we must set up a mutual fund which invests solely in the underlying asset in question with full reinvestment of dividends. This mutual fund is a tradeable asset.

interest rates stay positive because of the square-root of $r(t)$ in the volatility. Both are also examples of *affine term-structure models*: that is, the $D(t,T)$ can be written in the form

$$D(t,T) = \exp(A(T-t) - B(T-t)r(t))$$

for suitable functions A and B.

2.1 No-arbitrage models

The Vasicek and CIR models are examples of *time-homogeneous*, equilibrium models. A disadvantage of such models is that they give a set of theoretical prices for bonds which will not normally match precisely the actual prices that we observe in the market. This led to the development of some time-*inhomogeneous* Markov models for $r(t)$, most notably those due to Ho & Lee (1986)

$$dr(t) = \phi(t)dt + \sigma d\tilde{W}(t)$$

Hull & White (1990)

$$dr(t) = \alpha(\mu(t) - r(t))dt + \sigma(t)d\tilde{W}(t)$$

and Black & Karasinski (1991)

$$dr(t) = \alpha(t)r(t)\left[\theta(t) - \log r(t)\right]dt + \sigma(t)r(t)d\tilde{W}(t).$$

In each of these models $\phi(t)$, $\mu(t)$, $\alpha(t)$, $\theta(t)$ and $\sigma(t)$ are all deterministic functions of t. These determinstic functions are *calibrated* in a way which gives a precise match at the start date (say time 0) between theoretical and observed prices of zero-coupon bonds (Ho & Lee) and possibly also some derivative prices. For example, at-the-money interest-rate caplet prices could be used to derive the volatility function, $\sigma(t)$, in the Hull & White model.

Because these models involve an initial calibration of the model to observed prices there is no arbitrage opportunity at the outset. Consequently these models are often descibed as *no-arbitrage models*. In contrast, the time-homogeneous models described earlier tell us that if the prices were to evolve in a particular way then the dynamics will be *arbitrage free*.[9]

The Ho & Lee (1986) and Hull & White (1990) models are also examples of affine term-structure models. The Black & Karasinski (1991) model does not yield any analytical solutions, other than that $r(t)$ is log-normally distributed. However, the

[9] Thus, all the models we are considering here are *arbitrage free* and we reserve the use of the term *no-arbitrage model* to those where we have a precise match, at least initially, between theoretical and observed prices.

BK model is amenable to the development of straightforward and fast numerical methods for both calibration of parameters and calculation of prices.

It is standard market practice to recalibrate the parameters and time-dependent, deterministic functions in these no-arbitrage models on a frequent basis. For example, take two distinct times $T_1 < T_2$. In the Hull & White (1990) model we would calibrate at time T_1 the functions $\mu(t)$ and $\sigma(t)$ for all $t > T_1$ to market prices. Call these functions $\mu_1(t)$ and $\sigma_1(t)$. At time T_2 we would repeat this calibration using prices at T_2 resulting in functions $\mu_2(t)$ and $\sigma_2(t)$ for $t > T_2$. If the Hull & White model is correct then we should find that $\mu_1(t) = \mu_2(t)$ and $\sigma_1(t) = \sigma_2(t)$ for all $t > T_2$. In practice this rarely happens so that we end up treating $\mu(t)$ and $\sigma(t)$ as stochastic rather than the deterministic form assumed in the model. Users of this approach to calibration need to be (and are) aware of this inconsistency between the model assumptions and the approach to recalibration. Despite this drawback, practitioners do still calibrate models in this way so, one assumes (hopes?), that the impact of model error is not too great.

2.2 Multifactor models

The risk-neutral approach to pricing is easily extended (at least theoretically) to multifactor models. One approach models an n-dimensional diffusion process $X(t)$ with SDE

$$dX(t) = \mu(t, X(t))dt + \nu(t, X(t))d\tilde{W}(t)$$

where $\mu(t, X(t))$ is an $n \times 1$ vector, $\tilde{W}(t)$ is standard n-dimensional Brownian motion under the risk-neutral measure Q and $\nu(t, X(t))$ is the $n \times n$ matrix of volatilities. The risk-free rate of interest is then defined as a function g of $X(t)$: that is, $r(t) = g(X(t))$. Zero-coupon bond prices then have essentially the same form as before: that is,

$$D(t, T) = E_Q \left[\exp\left(-\int_t^T r(u)du \right) \;\middle|\; \mathcal{F}_t \right]$$

as do derivatives

$$V(t) = E_Q \left[\exp\left(-\int_t^T r(u)du \right) V(T) \;\middle|\; \mathcal{F}_t \right]. \tag{5}$$

However, in a Markov context, whereas the conditioning was on $r(t)$ in the one-factor model, we now have to condition on the whole of $X(t)$.

In some multifactor models the first component of $X(t)$ is equal to $r(t)$, but we need to remember that the future dynamics of $r(t)$ still depend upon the whole of $X(t)$. In other cases $r(t)$ is a linear combination of the $X_i(t)$ (for example, the Longstaff & Schwartz (1995) model). Brennan & Schwartz (1979) model

$X_1(t) = r(t)$ and $X_2(t) = l(t)$, the yield on irredeemable coupon bonds. Rebonato (1998, Chapter 15) models $X_1(t) = l(t)$ and $X_2(t) = r(t)/l(t)$, both as log-normal processes. However, both of the Brennan & Schwartz and Rebonato models are prone to instability and need to be used with great care.

3 The Heath-Jarrow-Morton (HJM) framework

The new framework proposed by Heath, Jarrow & Morton (1992) represented a substantial leap forward in how the term structure of interest rates is perceived and modelled. Previously models concentrated on modelling of $r(t)$ and other relevant quantities in a multifactor model. The HJM framework arose out of the question: if we look at, say, the Ho & Lee or the Hull & White models, how does the whole of the forward-rate curve evolve over time? So, instead of focusing on $r(t)$ and calculation of the expectation in equation (3), HJM developed a framework where we can model the instantaneous forward-rate curve, $f(t,T)$, directly. Given the forward-rate curve we then immediately get $D(t,T) = \exp\left[-\int_t^T f(t,u)du\right]$.

In the general framework we have the SDE

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)'dW(t) \tag{6}$$

for each fixed $T > t$, and where $\alpha(t,T)$ is the drift process (scalar), $\sigma(t,T)$ is an $n \times 1$ vector process and $W(t)$ is a standard n-dimensional Brownian motion under the real-world measure P. For such a model to be arbitrage free there must exist an $n \times 1$ vector process $\gamma(t)$ (the market prices of risk) such that

$$\begin{aligned} \alpha(t,T) &= \sigma(t,T)'\left(\gamma(t) - S(t,T)\right) \\ \text{where} \quad S(t,T) &= -\int_t^T \sigma(t,u)du. \end{aligned}$$

The vector process $S(t,T)$ is of interest in its own right as it gives us the volatilities of the zero-coupon bonds: that is,

$$\begin{aligned} dD(t,T) &= D(t,T)\left[\left(r(t) + S(t,T)'\gamma(t)\right)dt + S(t,T)'dW(t)\right] \\ \text{or} \quad dD(t,T) &= D(t,T)\left[r(t)dt + S(t,T)'d\tilde{W}(t)\right] \end{aligned} \tag{7}$$

where $\tilde{W}(t)$ is a Q-Brownian motion. The last equation is central to model-building under Q: unless we can write down the dynamics of a tradeable asset with $r(t)$ as the expected rate of return under Q we will not have an arbitrage-free model.

In fact equation (7) is often taken as the starting point in a pricing exercise built on the HJM framework. We derive the bond volatilities $S(t,T)$ from derivative prices and this then provides us with sufficient information to derive the prices of related derivatives. The market price of risk $\gamma(t)$ is only required if we are using the model in an asset-liability modelling or DFA exercise where real-world dynamics are required.

Returning to equation (6), the vector $\sigma(t, T)$ gives us what is described as the *volatility term-structure*. Developing this volatility function is central to model building and accurate pricing of derivatives under the HJM framework.

All of the short-rate models described earlier can be formulated within the HJM framework. For example, the Vasicek and Hull & White models both have $\sigma(t, T) = \sigma \exp(-\alpha(T - t))$, which is deterministic. In contrast, the CIR model has a stochastic volatility function which is proportional to $\sqrt{r(t)}$. However, the point of using the HJM framework is to think in much more general terms rather than restrict ourselves to the use of the earlier short-rate models. This added flexibility has made the approach popular with practitioners because it allows easy calibration: first, of the forward rates from, for example, the LIBOR term structure; second of the volatility term structure by making reference to suitable interest-rate derivatives.

Despite the advantages of HJM over the short-rate models it was found to have some some drawbacks: some practical, some theoretical. First, many volatility term structures $\sigma(t, T)$ result in dynamics for $f(t, T)$ which are non-Markov (that is, with a finite state space). This introduces path dependency to pricing problems which significantly increases computational times. Second there are generally no simple formulae or methods for pricing commonly-traded derivatives such as caps and swaptions. Again this is a significant problem from the computational point of view. Third, if we model forward rates as log-normal processes then the HJM model will "explode" (for example, see Sandmann & Sondermann, 1997). This last theoretical problem with the model can be avoided by modelling LIBOR and swap rates as log-normal (market models) rather than instantaneous forward rates.

4 Other assets as numeraires

Another essential step forward in term-structure modelling was the realisation that it is not necessary to use the cash account, $C(t)$, as the numeraire or the risk-neutral measure for pricing.

Instead we can follow the following guidelines.

- Let $X(t)$ be the price at time t of any tradeable asset which remains strictly positive at all times.

- Let $V(t)$ be the price of another tradeable asset.
- Find the measure Q_X equivalent to Q under which $V(t)/X(t)$ is a martingale. Here we find first the SDE for $V(t)/X(t)$ under Q. Thus given

$$\begin{aligned} dV(t) &= V(t)\left[r(t)dt + \sigma_V(t)'d\tilde{W}(t)\right] \\ \text{and } dX(t) &= X(t)\left[r(t)dt + \sigma_X(t)'d\tilde{W}(t)\right] \\ \text{we get } d\left(\frac{V(t)}{X(t)}\right) &= \frac{V(t)}{X(t)}\left(\sigma_V(t) - \sigma_X(t)\right)'\left(d\tilde{W}(t) - \sigma_X(t)dt\right). \end{aligned}$$

  Now define $W^X(t) = \tilde{W}(t) - \int_0^t \sigma_X(u)du$. We can then call on the Cameron-Martin-Girsanov Theorem (see, for example, Baxter & Rennie (1996)). This tells us that (provided $\sigma_X(t)$ satisfies the Novikov condition) there exists a measure Q_X equivalent to Q under which $W^X(t)$ is a standard n-dimensional Brownian motion. We then have

$$d\left(\frac{V(t)}{X(t)}\right) = \frac{V(t)}{X(t)}\left(\sigma_V(t) - \sigma_X(t)\right)' dW^X(t).$$

  In other words, $V(t)/X(t)$ is a martingale under Q_X as required.
- Now note that the transformation from Q to Q_X was independent of the choice of $V(t)$ so that the prices of all tradeable assets divided by the numeraire $X(t)$ are martingales under the same measure Q_X.

This has some important consequences. First, suppose that we have a derivative contract which pays $V(T)$ at time T. We can take $X(t) = D(t,T)$ as the numeraire. Rather than use the previous notation, Q_X, for the new measure, it is common to use Q_T for this specific type of numeraire. Q_T is called a *forward measure*. The martingale property tells us that

$$\begin{aligned} \frac{V(t)}{D(t,T)} &= E_{Q_T}\left[\left.\frac{V(T)}{D(T,T)}\right|\mathcal{F}_t\right] \\ \Rightarrow V(t) &= D(t,T)E_{Q_T}\left[V(T)|\mathcal{F}_t\right] \end{aligned}$$

since $D(T,T) = 1$. This pricing equation is certainly more appealing than the risk-neutral pricing equations (4) and (5) because we no longer need to work out the joint distribution of $V(T)$ and $\exp\left(-\int_t^T r(u)du\right)$ given $\mathcal{F}_t$. On the other hand we still need to establish the distribution of $V(T)$ under the new measure Q_T. How easy this is to do will depend upon the model being used.

For general numeraires $X(t)$ we also have the general derivative pricing equation

$$\frac{V(t)}{X(t)} = E_{Q_X}\left[\left.\frac{V(T)}{X(T)}\right|\mathcal{F}_t\right]. \tag{8}$$

Specific frameworks which make further use of this change of measure are the LIBOR market model (where we use $X(t)D(t,T_k)$ for derivative payments at T_k) and the swap market model (where we use $X(t) = \sum_{k=1}^{n} D(t,T_k)$ in relation to a swaption contract with exercise date T_0 and where $T_k = T_0 + k\tau$). Further details can be found in the accompanying article on *market models*.

5 State-price deflators and kernel functions

We have noted in the previous section that we can price derivatives under Q or Q_X. If we make a careful choice of our numeraire, $X(t)$, then we can turn the pricing problem back into one involving expectations under the real-world measure P.

The existence of P and its equivalence to the risk-neutral measure Q means that there exists a vector process $\gamma(t)$ which connects the risk-neutral and real-world Brownian motions: that is

$$d\tilde{W}(t) = dW(t) + \gamma(t)dt.$$

If we have a complete market then we can synthesize a self-financing portfolio with value $X(t)$ and SDE

$$\begin{aligned} dX(t) &= X(t)\left[r(t)dt + \gamma(t)'d\tilde{W}(t)\right] \quad \text{under } Q \\ &= X(t)\left[\left(r(t) + |\gamma(t)|^2\right)dt + \gamma(t)'dW(t)\right] \quad \text{under } P. \end{aligned}$$

Now define $A(t) = X(t)^{-1}$. Then the derivative pricing equation (8) becomes

$$V(t) = \frac{E_P[V(T)A(T)|\mathcal{F}_t]}{A(t)}.$$

The process $A(t)$ has a variety of different names: *state-price deflator*; *deflator*; and *pricing kernel*.

State-price deflators provide a useful theoretical framework when we are working in a multicurrency setting. In such a setting there is a different risk-neutral measure Q_i for each currency i. In contrast, the measure used in pricing with state-price deflators is unaffected by the base currency. Pricing using the risk-neutral approach is still no problem, but we just need to be a bit more careful.

Apart from this both the risk-neutral and state-price-deflator approaches to pricing have their pros and cons. One particular point to note is that specific models often have a natural description under one approach only. For example, the CIR model has a straightforward description under the risk-neutral approach, whereas attempting to describe it using the state-price-deflator approach is unnecessarily cumbersome and quite unhelpful. For other models the reverse is true: they have a relatively simple formulation using the state-price-deflator approach and are quite difficult to convert into the risk-neutral approach.

6 Positive interest

In this section we will describe what is sometimes referred to the *Flesaker-Hughston framework* (Flesaker & Hughston (1996)) or the *Potential approach* (Rogers (1997)).

This approach is a simple variation on the state-price-deflator approach and, with a seemingly simple constraint, we can guarantee that interest rates will remain positive (hence the name of this section).

The approach is deceptively simple. We start with a sample space Ω, with sigma-algebra $\mathcal{F}$ and associated filtration $\mathcal{F}_t$. Let $A(t)$ be a strictly positive diffusion process adapted to $\mathcal{F}_t$ and let $\hat{P}$ be some probability measure associated with $(\Omega, \mathcal{F}, \{\mathcal{F}_t : 0 \leq t < \infty\})$. $\hat{P}$ is called the *pricing measure* and may be different from the equivalent, real-world measure P.

Rogers (1997) and Rutkowski (1997) investigated the family of processes

$$D(t,T) = \frac{E_{\hat{P}}[A(T)|\mathcal{F}_t]}{A(t)} \quad \text{for } 0 \leq t \leq T$$

for all $T > 0$. They proved that the processes $D(t,T)$ can be regarded as the prices of zero-coupon bonds in the usual way, and that the proposed framework gives arbitrage-free dynamics. Additionally, if $V(T)$ is some $\mathcal{F}_T$-measureable derivative payoff at time T then the price for this derivative at some earlier time t is

$$V(t) = \frac{E_{\hat{P}}[A(T)V(T)|\mathcal{F}_t]}{A(t)}.$$

In addition, if $A(t)$ is a supermartingale under $\hat{P}$, (that is, $E_{\hat{P}}[A(T)|\mathcal{F}_t] \leq A(t)$ for all $0 < t < T$) then, for each t, $D(t,T)$ is a decreasing function of T implying that all interest rates are positive.

A year earlier Flesaker & Hughston (1996) proposed a special case of this where

$$A(t) = \int_t^\infty \phi(s) M(t,s) ds$$

for some family of strictly-positive $\hat{P}$-martingales $M(t,s)$.

Despite the apparent simplicity of the pricing formulae, all of the potential difficulties come in the model for $A(t)$ and in calculating its expectation. Specific examples were given in the original papers by Flesaker & Hughston (1996), Rutkowski (1997) and Rogers (1997). More recently Cairns (1999, 2003) used the Flesaker-Hughston framework to develop a family of models suitable for a variety of purposes from short-term derivative pricing to long-term risk management, such as Dynamic Financial Analysis.

7 Markov-functional models

Many (but not all) of the preceeding models can be described as *Markov-functional models*. This approach to modelling was first described by Hunt, Kennedy & Pelsser (2000) (HKP) with further accounts in Hunt & Kennedy (2000) and Pelsser (2000). In many of these models we were able to write down the prices of both bonds and derivatives as functions of a Markov process $X(t)$. For example, in most of the short-rate models (Vasicek, CIR, Ho-Lee, Hull-White) prices are all *functions* of the *Markov*, risk-free rate of interest, $r(t)$. Equally, with the positive-interest family of Cairns (1999, 2003) all prices are functions of a multi-dimensional Ornstein-Uhlenbeck process, $X(t)$, (which is Markov) under the pricing measure $\hat{P}$.

HKP generalise this as follows. Let $X(t)$ be some low-dimensional, time-homogeneous Markov diffusion process. Under a given pricing measure, $\hat{P}$, equivalent to the real-world measure P suppose that $X(t)$ is a martingale, and that prices are of the form

$$D(t, T) = f(t, T, X(t)).$$

Clearly, the form of the function f needs to be restricted to ensure that prices are arbitrage free. For example, we can employ the Potential approach to define the numeraire, $A(t)$, as a strictly-positive function of $X(t)$.

For a model to be a *Markov-functional model* HKP add the requirement that it should be possible to calculate relevant prices efficiently: for example, the prices of caplets and swaptions. Several useful examples can be found in Pelsser (2000) and Hunt & Kennedy (2000).

8 References

1. Baxter, M., and Rennie, A. (1996) *Financial Calculus.* CUP.

2. Black, F., and Karasinski, P. (1991) Bond and option pricing when short rates are log-normal. Financial Analysts Journal **July-Aug**, 52-29.

3. Black, F., and Scholes, M. (1973) The pricing of options and corporate liabilities. Journal of Political Economy **81**, 637-659.

4. Brennan, M., and Schwartz, E. (1979) A continuous-time approach to the pricing of bonds. Journal of Banking and Finance **3**, 133-155.

5. Brigo, D., and Mercurio, F. (2001) *Interest Rate Models: Theory and Practice.* Springer.

6. Cairns, A.J.G. (1998) Descriptive bond-yield and forward-rate models for the British government securities' market. British Actuarial Journal **4**, 265-321.

7. Cairns, A.J.G. (1999) A multifactor model for the term structure and inflation. In Proceedings of the 9th AFIR Colloquium, Tokyo, Volume 3: 93-113..

8. Cairns, A.J.G. (2003) A family of term-structure models for long-term risk management and derivative pricing. To appear in *Mathematical Finance.*.

9. Cairns, A.J.G. (2003) *Interest-Rate Models: An Introduction.* Princeton University Press.

10. Cairns, A.J.G., and Pritchard, D.J. (2001) Stability of descriptive models for the term-structure of interest rates with application to German market data. British Actuarial Journal **7**, 467-507.

11. Constantinides, G.M. (1992) A theory of the nominal term structure of interest rates. Review of Financial Studies **5**, 531-552.

12. Cox, J., Ingersoll, J., and Ross, S. (1985) A theory of the term-structure of interest rates. Econometrica **53**, 385-408.

13. Fisher, M., Nychka, D., and Zervos, D. (1995) Fitting the term structure of interest rates with smoothing splines. The Federal Reserve Board, Finance and Economics Discussion Series, Working Paper 1995-1..

14. Flesaker, B., and Hughston, L.P. (1996) Positive interest. Risk **9(1)**, 46-49.

15. Heath, D., Jarrow, R. and Morton, A. (1992) Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica **60**, 77-105.

16. Ho, S.Y., and Lee S-B. (1986) Term structure movements and pricing interest rate contingent claims. Journal of Finance **41**, 1011-1029.

17. Hull, J.C. (2002) *Options, Futures, and Other Derivatives, 5th Edition.* Prentice Hall.

18. Hull, J.C., and White, A.D. (1990) Pricing interest rate derivative securities. The Review of Financial Studies **3**, 573-592.

19. Hunt, P.J., and Kennedy, J.E. (2000) *Financial Derivatives in Theory and Practice.* Wiley.

20. Hunt, P.J., Kennedy, J.E., and Pelsser, A. (2000) Markov-functional interest-rate models. Finance and Stochastics **4**, 391-408.

21. James, J., and Webber, N. (2000) *Interest Rate Modelling.* Wiley.

22. Longstaff, F.A., and Schwartz, E.S. (1995) A simple approach to valuing risky fixed and floating-rate debt. Journal of Finance **50**, 789-819.

23. Merton, R.C. (1973) Theory of rational option pricing. Bell Journal of Economics and Management Science **4**, 141-183.

24. Nelson, C.R., and Siegel, A.F. (1987) Parsimonious modeling of yield curves. Journal of Business **60**, 473-489.

25. Pelsser, A. (2000) *Efficient Methods for Valuing Interest Rate Derivatives.* Springer.

26. Rebonato, R. (1998) *Interest Rate Option Models, Second Edition.* Wiley.

27. Rebonato, R. (2002) *Modern Pricing of Interest-Rate Derivatives: the LIBOR Market Model and Beyond.* Princeton University Press.

28. Rebonato, R. (2003) Term structure models: a review. Royal Bank of Scotland Quantitative Research Centre Working Paper.

29. Rogers, L.C.G. (1997) The potential approach to the term-structure of interest rates and foreign exchange rates. Mathematical Finance **7**, 157-164.

30. Rutkowski, M. (1997) A note on the Flesaker & Hughston model of the term structure of interest rates. Applied Mathematical Finance **4**, 151-163.

31. Sandmann, K., and Sondermann, D. (1997) A note on the stability of lognormal interest rate models and the pricing of Eurodollar futures. Mathematical Finance **7**, 119-128.

32. Schaefer, M.S., and Schwartz, E.S. (1984) A two-factor model of the term structure – an approximate analytical solution. Journal of Financial and Quantitative Analysis **19**, 413-424.

33. Svensson, L.E.O. (1994) Estimating and interpreting forward interest rates: Sweden 1992-1994. Working paper of the International Monetary Fund **94.114**, 33 pages.

34. Vasicek, O. (1977) An equilibrium characterisation of the term structure. Journal of Financial Economics **5**, 177-188.

35. Wilkie, A.D. (1995) More on a stochastic asset model for actuarial use (with discussion). British Actuarial Journal **1**, 777-964.