# Entropy and Information in the Interest Rate Term Structure

## D. C. Brody[*] and L. P. Hughston[†]

[*]Blackett Laboratory, Imperial College, London SW7 2BZ, UK
[†]Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK

**Abstract.** Associated with every positive interest term structure there is a probability density function over the positive half line. This fact can be used to turn the problem of term structure analysis into a problem in the comparison of probability distributions, an area well developed in statistics, known as information geometry. The information-theoretic and geometric aspects of term structures thus arising are here illustrated. In particular, we introduce a new term structure calibration methodology based on maximisation of entropy, and also present some new families of interest rate models arising naturally in this context.

## 1. Introduction

Dynamical models for interest rates suffer from the fact that it is difficult to isolate the independent degrees of freedom in the evolution of the term structure. The question is, which ingredients in the determination of an interest rate model can and should be specified independently and exogenously? A related issue, important for applications, is the determination of an appropriate data set for the specification of initial conditions—the so-called 'calibration problem'.

Traditionally interest rate models have tended to focus either on discount bonds or on rates. Depending on which choice is made, the resulting models take different forms, and hence have a different feel to them. Fundamentally, however, it should not make any difference whether a model is based on bonds or rates. To develop this point further we introduce some notation. Let time 0 denote the present. We write $P_{tT}$ for the value at time $t$ of a discount bond that matures to deliver one unit of currency at time $T$. The associated interest rate $L_{tT}$, expressed on a simple basis, is determined by the relation

$$P_{tT} = \frac{1}{1 + (T - t)L_{tT}}. \tag{1}$$

Alternatively, we may use the associated continuously compounded rate $R_{tT}$ defined by

$$P_{tT} = \exp\left(-(T-t)R_{tT}\right). \tag{2}$$

The dynamics of $L_{tT}$, $R_{tT}$, and $P_{tT}$ all look different from one another, even if the underlying model is the same.

Now suppose we write $P_{tab}$ for the forward price, contracted at time $t$, for delivery of a discount bond at time $a$ that matures at time $b$. Then a standard arbitrage argument shows that

$$P_{tab} = \frac{P_{tb}}{P_{ta}}. \tag{3}$$

The corresponding forward interest $F_{tab}$, contracted at time $t$ for the period beginning at time $a$ and ending at time $b$, expressed on a simple basis, is defined by the relation

$$P_{tab} = \frac{1}{1+(b-a)F_{tab}}. \tag{4}$$

The associated continuously compounded forward rate $R_{tab}$ is defined analogously by the relation

$$P_{tab} = \exp\left(-(b-a)R_{tab}\right). \tag{5}$$

Two other related interest rates of importance are the short rate

$$r_t = L_{tt} \tag{6}$$

and the forward short rate

$$f_{tT} = F_{tTT} \tag{7}$$

also known as the instantaneous forward rate. These rates are defined by applying limiting arguments to (1) and (4). Alternatively, we can write

$$f_{tT} = -\frac{\partial}{\partial T}\ln P_{tT} \tag{8}$$

and

$$r_t = f_{tt}. \tag{9}$$

Once we have the discount bond system, the associated rates can be directly constructed. It follows that it is sufficient to consider the problem of examining the independent dynamical degrees of freedom in the discount bond system. We shall consider briefly two examples of useful ways in which this can be done for general interest rate models, indicating as well the associated drawbacks.

*Example 1. Dynamic models for the short rate.* In this case the model is defined with respect to a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\{\mathcal{F}_t\}$ and a standard multidimensional Brownian motion $W_t^\alpha$ ($\alpha = 1, 2, \cdots, n$), where $n$ is possibly infinite. The independent degrees of freedom are given by (a) the specification of the short rate $r_t$ as an essentially arbitrary Ito process on $(\Omega, \mathcal{F}, \mathbb{P})$, and (b) a market risk premium process $\lambda_t^\alpha$. The model for the discount bonds is

$$P_{tT} = \frac{1}{\Lambda_t}\mathbb{E}_t\left[\Lambda_T \exp\left(-\int_t^T r_s ds\right)\right], \tag{10}$$

where $\mathbb{E}_t$ denotes conditional expectation with respect to the filtration $\mathcal{F}_t$ and the density martingale $\Lambda_t$ is defined by

$$\Lambda_t = \exp\left(-\int_0^t \lambda_s dW_s - \tfrac{1}{2}\int_0^t \lambda_s^2 ds\right). \tag{11}$$

Here and in what follows we use the shortened notation

$$\lambda_s dW_s = \sum_{\alpha=1}^n \lambda_s^\alpha dW_s^\alpha \tag{12}$$

for the vector inner product. A similar convention is used for products such as $\lambda_s^2 = \sum_\alpha \lambda_s^\alpha \lambda_s^\alpha$. An advantage of this general model is that the processes $r_t$ and $\lambda_t^\alpha$ can be specified independently and exogenously, and for interest rate positivity it suffices to let the process $r_t$ be positive. There are two disadvantages to this approach. Firstly, the model is specified implicitly: the conditional expectation is generally difficult to calculate. Secondly, the initial term structure is not fed in directly.

A further simplification can be achieved in Example 1 by introducing the state price density process $Z_t$, given by

$$Z_t = \Lambda_t \exp\left(-\int_0^t r_s ds\right), \tag{13}$$

from which it follows that

$$P_{tT} = \frac{\mathbb{E}_t[Z_T]}{Z_t}. \tag{14}$$

Then it is sufficient to specify the pricing kernel $Z_t$ alone, and we can recover $r_t$ and $\lambda_t^\alpha$ from the relation

$$\frac{dZ_t}{Z_t} = -r_t dt - \lambda_t dW_t. \tag{15}$$

*Example 2. The Heath-Jarrow-Morton framework.* In this case the independent dynamical degrees of freedom consist of (a) the initial term structure $P_{0T}$, (b) the market risk premium process $\lambda_t^\alpha$, and (c) the forward short rate volatility process $\sigma_{tT}^\alpha$ for each maturity $T$. The model for the discount bonds is

$$P_{tT} = \exp\left(-\int_t^T f_{ts} ds\right). \tag{16}$$

Here the forward short rates are given by

$$f_{tT} = -\frac{\partial}{\partial T}\ln P_{0T} - \int_{s=0}^t \sigma_{sT}\Omega_{sT} ds + \int_{s=0}^t \sigma_{sT}(dW_s + \lambda_s ds), \tag{17}$$

where

$$\Omega_{tT}^\alpha = -\int_{u=t}^T \sigma_{tu}^\alpha du. \tag{18}$$

The advantage of the HJM framework is that it allows a direct input of the initial term structure, as well as control over the volatility structure of the discount bonds. Indeed, it follows from (16), (17), and (18) that

$$\frac{dP_{tT}}{P_{tT}} = r_t dt + \Omega_{tT}(dW_t + \lambda_t dt), \tag{19}$$

where $r_t$ is given by (10), and as a consequence of (18) we see that the discount bond volatility $\Omega_{tT}$ goes to zero in the limit $t \to T$ as the bond matures. A disadvantage of the HJM approach is that there is no guarantee of interest rate positivity, and it is not easy to impose a direct condition on $\sigma_{tT}^{\alpha}$ to achieve this.

In this article we explore an alternative framework for isolating the independent degrees of freedom in interest rate dynamics. Our approach has the virtue of retaining the desirable features of both examples cited above, while eliminating the undesirable features. The key idea, as we discuss in greater detail shortly, is the introduction of a term structure density process $\rho_t(x)$ defined by

$$\rho_t(x) = -\frac{\partial}{\partial x} B_{tx}. \tag{20}$$

Here $B_{tx}$ denotes the system of bond prices at time $t$ when we parameterise the bonds by the tenor variable $x = T - t$, so

$$B_{tx} = P_{t,t+x}, \tag{21}$$

and we make the reasonable assumption that $B_{tx} \to 0$ for large $x$. It is then a straightforward exercise to verify that the interest rate positivity conditions

$$0 < B_{tx} \leq 1, \quad \text{and} \quad \frac{\partial}{\partial x} B_{tx} < 0 \tag{22}$$

are equivalent to the following relations on $\rho_t(x)$:

$$\rho_t(x) > 0, \quad \text{and} \quad \int_0^\infty \rho_t(x) dx = 1. \tag{23}$$

We therefore reach the interesting conclusion that any positive interest rate model can be regarded as a random process on the space of density functions on the positive real line.

This observation ties in naturally with the general growing interest in modelling the interest rate term structure as a dynamical system (Björk & Christensen 1999, Björk & Gombani 1999, Björk 2000, Brody & Hughston 2001a,b, Filipović 2001). The idea is that we treat the yield curve as a mathematical object in its own right, identified as a "point" $\rho$ lying in the space $\mathfrak{M}$ of all possible yield curves. With the specification of an initial yield curve $\rho_0$ we model the resulting dynamics as a random trajectory $\rho_t$ in $\mathfrak{M}$. By bringing the structure of $\mathfrak{M}$ into play it is possible both to clarify the status of existing interest models, and also to devise new interest rate models. In what follows we shall sketch some examples of new models that can be developed by following this line of argument.

## 2. Admissible term structures

An interesting consequence of the characterisation of the positive interest property in terms of a density function is that there is a natural 'information geometry' associated with the space of yield curves, which we now proceed to describe. For this purpose it suffices for the moment simply to consider properties of the *initial* term structure. Let

$t = 0$ denote the present, and $P_{0x}$ a family of discount bond prices satisfying $P_{00} = 1$, where $x$ is the tenor $(0 \leq x < \infty)$. We impose the condition that interest rates should always be positive with the following criterion:

**Definition**. *A term structure is said to be admissible if the discount function $P_{0x}$ is of class $C^\infty$ and satisfies $0 < P_{0x} \leq 1$, $\partial_x P_{0x} < 0$, and $\lim_{x\to\infty} P_{0x} = 0$.*

An admissible discount function can be viewed as a complementary probability distribution. In other words, we can think of the tenor date as an abstract random variable $X$, and for its distribution write

$$\Pr[X < x] = 1 - P_{0x}. \tag{24}$$

The associated density function $\rho(x) = -\partial_x P_{0x}$ satisfies $\rho(x) > 0$ for all $x$, and $\int_x^\infty \rho(u)du = P_{0x}$. We say that a density function is *smooth* if it is of class $C^\infty$ on the positive half-line $\mathbb{R}_+^1 = [0,\infty)$. Then we have the following characterisation of the space $\mathfrak{M}$:
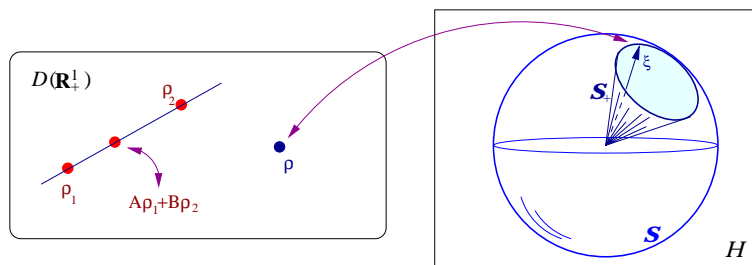
**Proposition 1**. *The system of admissible term structures is isomorphic to the convex space $\mathcal{D}(\mathbb{R}_+^1)$ of everywhere positive smooth density functions on the positive real line.*

This idea is illustrated in Figure 1. The requirement that $P_{0x}$ should be of class $C^\infty$ can be weakened, but we would argue that in practice any term structure can be approximated arbitrarily closely by a 'nearby' term structure with a smooth density. It is reasonable at least to insist that the forward short rate curve $f_{0x} = -\partial_x \ln P_{0x}$ is piecewise continuous and nonvanishing for all $x < \infty$.

Given a pair of term structure densities $\rho_1(x)$ and $\rho_2(x)$ we can define a natural distance function $\phi_{12}$ on $\mathfrak{M}$ by

$$\phi_{12} = \cos^{-1} \int_0^\infty \xi_1(x)\xi_2(x)dx, \tag{25}$$

where $\xi_i(x) = \sqrt{\rho_i(x)}$. We call this angle the Bhattacharyya distance between the given yield curves (cf. Bhattacharyya 1943). The geometrical interpretation of $\phi_{12}$ arises from



**Figure 1.** *The system of admissible term structures.* A smooth positive interest term structure can be regarded as a point in $\mathcal{D}(\mathbb{R}_+^1)$, the convex space of smooth and everywhere positive density functions on the positive half-line $\mathbb{R}_+^1$. Associated with each point $\rho \in \mathcal{D}(\mathbb{R}_+^1)$ there is a ray $\xi$ lying in the positive orthant $\mathcal{S}_+$ of the unit sphere $\mathcal{S}$ in the Hilbert space $\mathcal{H} = L^2(\mathbb{R}_+^1)$. A dynamical trajectory on $\mathcal{D}(\mathbb{R}_+^1)$ can then be mapped to a corresponding trajectory in $\mathcal{S}_+$.

the fact that the map $\rho(x) \to \xi(x)$ associates to each point of $\mathfrak{M}$ a point in the positive orthant $\mathcal{S}^+$ of the unit sphere in the Hilbert space $L^2(\mathbb{R}_+^1)$, and $\phi_{12}$ is the resulting spherical angle on $\mathcal{S}^+$. Note that $0 \leq \phi < \frac{1}{2}\pi$ and that orthogonality can never be achieved if forward rates are nonvanishing.

As a simple illustration we consider the family of discount bonds given by

$$P_{0x} = \left(1 + \frac{Rx}{\kappa}\right)^{-\kappa}, \tag{26}$$

where $R$ and $\kappa$ are constants. In this case we have a flat term structure, with a constant annualised rate of interest $R$ assuming compounding at the frequency $\kappa$ over the life of each bond ($\kappa$ need not be an integer). For $\kappa = 1$ this reduces to the case of a flat rate on the basis of a simple yield, and in the limit $\kappa \to \infty$ we recover the case of a flat rate on the basis of continuous compounding. For the density function $\rho(x) = -\partial_x P_{0x}$ associated with (26) we obtain

$$\rho(x) = R\left(1 + \frac{Rx}{\kappa}\right)^{-(\kappa+1)}. \tag{27}$$

Let us write $\rho_i(x)$ for the density corresponding to $R = R_i$ ($i = 1, 2$) for a fixed value of $\kappa$. A direct calculation of the integral (25) gives

$$\phi_{12} = \cos^{-1}\left(\frac{\sqrt{R_1 R_2}}{R_1 - R_2} \log \frac{R_1}{R_2}\right) \tag{28}$$

for the distance when $\kappa = 1$. In the limit $\kappa \to \infty$ (continuous compounding) we have

$$\phi_{12} = \cos^{-1}\left(\frac{2\sqrt{R_1 R_2}}{R_1 + R_2}\right), \tag{29}$$

where the bracketed term in (29) is the ratio of the geometric and arithmetic means of the two rates. Note that in this limit we have $\rho(x) \to Re^{-Rx}$.

## 3. Dynamics of the term structure density

Now let us consider in more detail the evolution of the term structure density. Here we follow the line of argument presented in Brody & Hughston (2001a,b). We write $P_{tT}$ for the random value at time $t$ of a discount bond that matures at time $T$, where $T \in \mathbb{R}_+^1$ and $0 \leq t \leq T$, and assume, for each $T$, that $P_{tT}$ is an Ito process on the interval $t \in [0, T]$, for which the dynamics can be expressed in the form

$$dP_{tT} = m_{tT}dt + \Sigma_{tT}dW_t. \tag{30}$$

The bond price process $P_{tT}$ is assumed for each value of $T \geq t$ to be adapted to the augmented filtration generated by $W_t$. The absolute drift $m_{tT}$ and the absolute volatility process $\Sigma_{tT}$ are assumed to satisfy regularity conditions sufficient to ensure that $\partial_T P_{tT}$ is also an Ito process. For interest rate positivity we require $0 < P_{tT} \leq 1$ and $\partial_T P_{tT} < 0$. Additionally we impose the asymptotic conditions $\lim_{T\to\infty} P_{tT} = 0$, and $\lim_{T\to\infty} \partial_T P_{tT} = 0$.

We define the forward short rate $f_{tT}$ and the short rate $r_t$ as in (8) and (9). Because $P_{tT}$ is positive, $f_{tT}$ is an Ito process iff $-\partial_T P_{tT}$ is an Ito process. Note that $f_{tT}$ is the forward rate fixed at time $t$ for the short rate at time $T$.

For no arbitrage we require the existence of an exogenous market risk premium process $\lambda_t$ such that the absolute drift is of the form

$$m_{tT} = r_t P_{tT} + \lambda_t \Sigma_{tT}. \tag{31}$$

We do not assume the bond market is complete. If the bond market is complete, however, then $\lambda_t$ is determined endogenously by the bond price system.

We introduce the Musiela parameterisation $x = T - t$ for the tenor, and write $B_{tx} = P_{t,t+x}$ for the price at time $t$ of a bond for which the time to maturity is $x$. As a consequence of (30) and (31) we have the following dynamics for $B_{tx}$:

$$dB_{tx} = (r_t - f_{t,t+x})B_{tx}dt + \Sigma_{t,t+x}(dW_t + \lambda_t dt). \tag{32}$$

Now suppose we consider the time dependent term structure density $\rho_t(x)$ defined by (20), for which we have the normalisation condition

$$\int_{x=0}^{\infty} \rho_t(x)dx = 1, \tag{33}$$

or equivalently

$$\int_{u=t}^{\infty} \rho_t(u - t)du = 1. \tag{34}$$

The relation

$$\rho_t(T - t) = f_{tT}P_{tT} \tag{35}$$

then allows us to deduce an interesting interpretation of the normalisation condition. In particular, the formula

$$\int_t^{\infty} P_{tu}f_{tu}du = 1 \tag{36}$$

says that the value at time $t$ of a continuous cash flow in perpetuity that pays the small amount $f_{tu}du$ at time $u$ is always unity. Thus we can think of $f_{tu}$ as defining the 'convenience yield' associated with a position in cash. An analogous calculation shows that

$$\int_t^{\infty} P_{tu}^{\kappa}f_{tu}du = \frac{1}{\kappa} \tag{37}$$

for any positive value of the exponent $\kappa$. This relation can be interpreted by saying that if we 'fix' the convenience yield (e.g., by swapping the unit of cash for the corresponding future cash flow), and then rescale all the interest rates $R_{tu}$ by the same factor $\kappa$, so $R_{tu} \to \kappa R_{tu}$ for all $u \geq t$, where $R_{tu}$ is given by (2), then the value of the promised cash flow scales inversely with respect to $\kappa$.

Returning now to the evolutionary equation (32), suppose we write

$$\omega_{tx} = -\partial_x \Sigma_{t,t+x}. \tag{38}$$

Then we obtain the following dynamics for $\rho_t(x)$:

$$d\rho_t(x) = (r_t\rho_t(x) + \partial_x\rho_t(x))dt + \omega_{tx}(dW_t + \lambda_t dt). \tag{39}$$

The process $\omega_{tx}$ is subject to the constraint $\int_0^\infty \omega_{tx}dx = 0$, which implies that $\omega_{tx}$ is of the form

$$\omega_{tx} = \rho_t(x)(\nu_t(x) - \bar{\nu}_t), \tag{40}$$

where $\nu_t(x)$ is unconstrained, and

$$\bar{\nu}_t = \int_0^\infty \rho_t(u)\nu_t(u)du. \tag{41}$$

It follows from equation (38) that the absolute discount bond volatility $\Sigma_{tT}$ is given in the Musiela parameterisation by

$$\begin{aligned}
\Sigma_{t,t+x} &= \int_{u=x}^\infty \omega_{tu}du \\
&= \int_{u=x}^\infty \rho_t(u)\nu_t(u)du - \bar{\nu}_t B_{tx}.
\end{aligned} \tag{42}$$

This relation has an interesting probabilistic interpretation. Suppose, in particular, we write $I_x(u)$ for the indicator function

$$I_x(u) = \chi(u \geq x), \tag{43}$$

where $\chi(A)$ is unity if $A$ is true and vanishes otherwise. Then the bond price $B_{tx}$ can be written in the form of an abstract 'expectation':

$$\begin{aligned}
B_{tx} &= \int_{u=0}^\infty \rho_t(u)I_x(u)du \\
&= \mathbb{M}_t\left[I_x\right],
\end{aligned} \tag{44}$$

where

$$\mathbb{M}_t[g] = \int_{u=0}^\infty \rho_t(u)g(u)du \tag{45}$$

for any function $g(x)$. The absolute discount bond volatility can then be expressed as an abstract covariance of the form

$$\Sigma_{t,t+x} = \mathbb{M}_t\left[I_x\nu_t\right] - \mathbb{M}_t\left[I_x\right]\mathbb{M}_t[\nu_t]. \tag{46}$$

As a consequence of (46) we see that the bond volatility structure $\Sigma_{t,t+x}$ is invariant under the transformation $\nu_t(x) \to \nu_t(x) + \alpha_t$, where $\alpha_t$ is independent of $x$. This also follows directly from (40). This 'gauge' freedom can be used to set $\lambda_t = -\bar{\nu}_t$. Then $\lambda_t$ and $\Sigma_{tx}$ are both determined by $\nu_t(x)$.

**Proposition 2**. *The general admissible term structure evolution based on the filtration generated by a Brownian motion $W_t$ on $\mathcal{H}$ is a measure valued process $\rho_t(x)$ on $\mathcal{D}(\mathbb{R}_+^1)$ that satisfies*

$$d\rho_t(x) = (r_t\rho_t(x) + \partial_x\rho_t(x))\,dt + \rho_t(x)\left(\nu_t(x) - \bar{\nu}_t\right)\left(dW_t - \bar{\nu}_t dt\right), \tag{47}$$

where $\bar{\nu}_t = \int_0^\infty \rho_t(u)\nu_t(u)du$. *The volatility structure $\nu_t(x)$ can be specified exogenously along with the initial term structure density $\rho_0(x)$. The associated short rate process $r_t = \rho_t(0)$ satisfies*

$$dr_t = \left(r_t^2 + \partial_x\rho_t(x)|_{x=0}\right)dt + r_t(\nu_t(0) - \bar{\nu}_t)(dW_t - \bar{\nu}_t dt). \tag{48}$$

**Proposition 3**. *The solution of the dynamical equation for $\rho_t(x)$ in terms of the volatility structure $\nu_t(x)$ and the initial term structure density $\rho_0(x)$ is*

$$\rho_t(T - t) = \rho_0(T)\frac{\exp\left(\int_{s=0}^t V_{sT}dW_s - \frac{1}{2}\int_{s=0}^t V_{sT}^2 ds\right)}{\int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}, \tag{49}$$

*where*

$$V_{tu} = \nu_t(u - t). \tag{50}$$

*Proof.* The second term in the drift on the right of (47) can be eliminated by setting $x = T - t$, which gives us

$$d\rho_t(T - t) = r_t\rho_t(T - t)dt + \rho_t(T - t)\left(\nu_t(T - t) - \bar{\nu}_t\right)(dW_t - \bar{\nu}_t dt). \tag{51}$$

Integrating this relation and separating out the terms involving $\bar{\nu}_t$ we obtain

$$\rho_t(T - t) = \rho_0(T)\frac{\exp\left(\int_{s=0}^t r_s ds + \int_{s=0}^t \nu_s(T - s)dW_s - \frac{1}{2}\int_{s=0}^t \nu_s^2(T - s)ds\right)}{\exp\left(\int_{s=0}^t \bar{\nu}_s dW_s - \frac{1}{2}\int_{s=0}^t \bar{\nu}_s^2 ds\right)}. \tag{52}$$

It follows by use of the definition (50) that

$$\rho_t(T - t) = \rho_0(T)\frac{\exp\left(\int_{s=0}^t r_s ds + \int_{s=0}^t V_{sT}dW_s - \frac{1}{2}\int_{s=0}^t V_{sT}^2 ds\right)}{\exp\left(\int_{s=0}^t \bar{\nu}_s dW_s - \frac{1}{2}\int_{s=0}^t \bar{\nu}_s^2 ds\right)}. \tag{53}$$

Then with an application of the normalisation condition (34) we deduce as a consequence of (53) that

$$\exp\left(-\int_{s=0}^t r_s ds + \int_{s=0}^t \bar{\nu}_s dW_s - \frac{1}{2}\int_{s=0}^t \bar{\nu}_s^2 ds\right)$$
$$= \int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du. \tag{54}$$

When this relation is inserted in the denominator of (53), we immediately obtain (49). $\diamond$

It is interesting in this connection to note, by setting $T = t$ in (49), that the short rate process is given by

$$r_t = \rho_0(t)\frac{\exp\left(\int_{s=0}^t V_{st}dW_s - \frac{1}{2}\int_{s=0}^t V_{st}^2 ds\right)}{\int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}. \tag{55}$$

We observe, in particular, that in a deterministic model, with $V_{st} = 0$, this formula reduces to $r_t = \rho_0(t)/\int_t^\infty \rho_0(u)du$, or, in other words, $r_t = f_{0t}$. For the market risk premium process it follows from (41) together with the relation $\lambda_t = -\bar{\nu}_t$ that

$$\lambda_t^\alpha = -\frac{\int_{u=t}^\infty \rho_0(u)V_{tu}^\alpha \exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}{\int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}. \tag{56}$$

These formulae show that, given the initial term structure density $\rho_0(x)$ and the volatility structure $\nu_t(x)$, we can reconstruct the short rate process and the market risk premium processes. We deduce from (49) that the corresponding formula for the bond price process is

$$P_{tT} = \frac{\int_{u=T}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}{\int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}. \tag{57}$$

For the unit-initialised money market account $B_t$, satisfying $dB_t = r_t B_t dt$ and $B_0 = 1$, we have

$$B_t = \frac{\exp\left(\int_{s=0}^t \bar{\nu}_s dW_s - \frac{1}{2}\int_{s=0}^t \bar{\nu}_s^2 ds\right)}{\int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}, \tag{58}$$

which follows directly from (54). The density martingale $\Lambda_t$ is given by

$$\Lambda_t = \exp\left(\int_{s=0}^t \bar{\nu}_s dW_s - \frac{1}{2}\int_{s=0}^t \bar{\nu}_s^2 ds\right), \tag{59}$$

and for the state price density we have

$$Z_t = \int_{u=t}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du. \tag{60}$$

As a consequence we can then check that $Z_t = \Lambda_t/B_t$.

If we divide (49) by (57) we are led to a recipe for constructing the general positive interest HJM forward short rate system $f_{tT}$ in terms of freely specified data. In particular, we obtain

$$f_{tT} = \rho_0(T)\frac{\exp\left(\int_{s=0}^t V_{sT}dW_s - \frac{1}{2}\int_{s=0}^t V_{sT}^2 ds\right)}{\int_{u=T}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}. \tag{61}$$

Note that when $T = t$ this expression reduces to formula (55). A short calculation then allows us to deduce the following result:

**Proposition 4**. *The general positive interest HJM forward short rate volatility structure is*

$$\sigma_{tT} = f_{tT}(V_{tT} - U_{tT}) \tag{62}$$

*where $f_{tT}$ is given by (61), and*

$$U_{tT} = \frac{\int_{u=T}^\infty \rho_0(u)V_{tu}\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}{\int_{u=T}^\infty \rho_0(u)\exp\left(\int_{s=0}^t V_{su}dW_s - \frac{1}{2}\int_{s=0}^t V_{su}^2 ds\right)du}. \tag{63}$$

*The initial term structure density $\rho_0(x)$ and the volatility structure $V_{tu}$ $(u \geq t)$ are freely specifiable.*

In other words, in the HJM theory the forward short rate volatility is not freely specifiable if the interest rates are to be positive. Instead it must be of the form (62) where $V_{tT}$ is freely specifiable, along with the initial term structure. This result establishes a connection between the present approach and Example 2, and takes a significant step towards the resolution of the outstanding difficulty associated with that example.

## 4. Construction of admissible models

As a consequence of Proposition 3 we see that the general term structure density can also be expressed in the form

$$\rho_t(x) = \frac{\rho_0(t+x)M_{t,t+x}}{\int_{u=0}^{\infty}\rho_0(t+u)M_{t,t+u}du}, \tag{64}$$

or equivalently

$$\rho_t(T-t) = \frac{\rho_0(T)M_{tT}}{\int_{u=t}^{\infty}\rho_0(u)M_{tu}du}, \tag{65}$$

where for each $T$ the process $M_{tT}$ is a martingale ($0 \leq t \leq T < \infty$) such that $M_{tT} > 0$ and $M_{0T} = 1$. The process $M_{tT}$ is the exponential martingale associated with $V_{tT}$. This expression for $\rho_t(T-t)$ arises also in the model of Flesaker and Hughston (1996), in which the discount bond system has the representation

$$P_{tT} = \frac{\int_{u=T}^{\infty}\rho_0(u)M_{tu}du}{\int_{u=t}^{\infty}\rho_0(u)M_{tu}du}. \tag{66}$$

For details, see, for example, Flesaker & Hughston (1997, 1998), Rutkowski (1997), Musiela & Rutkowski (1997), Rogers (1997), Brody (2000) and Hunt & Kennedy (2000).

*Quasi-lognormal models.* An interesting class of specific models is obtained if we restrict the Brownian motion to be one-dimensional and let the volatility structure $V_{tu} = \nu_t(u - t)$ appearing in (49) be *deterministic*. Then $V_{tu}$ is a function of two variables defined on the region $0 \leq t \leq u < \infty$. The resulting term structure model has a good deal of tractability and exhibits some desirable features. In particular, the function $V_{tu}$ has the right structure for allowing a calibration of the model to a family of implied caplet volatilities for a fixed strike (e.g., at-the-money). If the dimensionality of the Brownian motion is increased then other strikes can be incorporated as well.

*Semi-linear models.* Another interesting special case can be obtained if we write

$$\rho_0(u) = \int_0^{\infty} e^{-uR}\phi(R)dR, \tag{67}$$

for the initial term structure density, where $\phi(R)$ is the inverse Laplace transform of $\rho_0(u)$. Then for certain choices of the martingale family $M_{tu}$ the integration in (66) can

be carried out explicitly. An example can be obtained as follows. Let $M_t$ be a martingale $(0 \leq t < \infty)$ and $Q_t$ the associated quadratic variation satisfying $(dM_t)^2 = dQ_t$, and set

$$M_{tT} = \exp\left((\alpha + \beta T)M_t - \tfrac{1}{2}(\alpha + \beta T)^2 Q_t\right). \tag{68}$$

This model arises if we put

$$\nu_t(T - t) = (\alpha + \beta T)\sigma_t \tag{69}$$

in Proposition 2, where the process $\sigma_t$ is defined by $dM_t = \sigma_t dW_t$. Then the $u$-integration can be carried out explicitly in the expressions for $\rho_t(x)$ and $P_{tT}$, and the results can be expressed in closed form. In particular, the bond prices can be expressed in the form

$$P_{tT} = \frac{\int_{R=0}^{\infty} \phi(R) \left(\int_{u=T}^{\infty} \mathrm{e}^{-uR} M_{tu} du\right) dR}{\int_{R=0}^{\infty} \phi(R) \left(\int_{u=t}^{\infty} \mathrm{e}^{-uR} M_{tu} du\right) dR}. \tag{70}$$

Here the bracketed expression in the integrand in the numerator is given by:

$$\int_{u=T}^{\infty} \mathrm{e}^{-uR} M_{tu} du = \frac{1}{|\beta|\sqrt{Q_t}} \exp\left(\frac{(M_t - R/\beta)^2}{2Q_t} + \alpha R/\beta\right)$$
$$\times \mathcal{N}\left(\pm\frac{M_t - R/\beta}{\sqrt{Q_t}} \mp (\alpha + \beta T)\sqrt{Q_t}\right), \tag{71}$$

where

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\tfrac{1}{2}\xi^2\right) d\xi \tag{72}$$

is the normal distribution function, and the $\pm$ sign is chosen in accordance with the sign of $\beta$. For example, in the case of an initial term structure with a constant continuously compounding rate $r$, corresponding to the choice $\phi(R) = \delta(R - r)$, we obtain

$$P_{tT} = \frac{\mathcal{N}\left(\pm\frac{M_t - r/\beta}{\sqrt{Q_t}} \mp (\alpha + \beta T)\sqrt{Q_t}\right)}{\mathcal{N}\left(\pm\frac{M_t - r/\beta}{\sqrt{Q_t}} \mp (\alpha + \beta t)\sqrt{Q_t}\right)}. \tag{73}$$

## 5. Moment analysis and the role of the perpetual annuity

Some interesting aspects of the term structure dynamics are captured in the properties of the moments of $\rho_t(x)$, defined by

$$\bar{x}_t = \int_0^{\infty} x \rho_t(x) dx, \quad \bar{x}_t^{(n)} = \int_0^{\infty} (x - \bar{x}_t)^n \rho_t(x) dx \tag{74}$$

where $n \geq 2$. In some cases the relevant moments may not exist. For example, in the case of a continuously compounded flat yield curve given at $t = 0$ by the density function $\rho_0(x) = R\mathrm{e}^{-Rx}$, we have $\bar{x}_0 = R^{-1}$, $\bar{x}_0^{(2)} = R^{-2}$, $\bar{x}_0^{(3)} = 3R^{-3}$ and $\bar{x}_0^{(4)} = 9R^{-4}$, whereas in the case of the simple flat term structure, for which $\rho_0(x) = R/(1 + Rx)^2$, none of the moments exist. The first four moments, if they exist, are the mean, variance, skewness and kurtosis of the distribution of the 'abstract' random variable $X$ characterising the term structure. The mean $\bar{x}_t$ is a characteristic time-scale associated with the yield curve, and its inverse $1/\bar{x}_t$ is an associated characteristic yield. The financial significance of $\bar{x}_t$ will be discussed shortly.

Let us examine the mean and the variance processes of a general admissible arbitrage-free term structure. For simplicity we introduce the following notation for the variance process:

$$v_t = \int_0^\infty x^2 \rho_t(x) dx - (\bar{x}_t)^2. \tag{75}$$

We assume that $\rho_t(x)$ and the discount bond volatility $\Sigma_{t,t+x}$ fall off sufficiently rapidly to ensure that $\lim_{x\to\infty} x^n \rho_t(x) = 0$ and $\lim_{x\to\infty} x^n \Sigma_{t,t+x} = 0$ for $n = 1, 2$, and that the integrals $\int_0^\infty x^n \rho_t(x) dx$ and $\int_0^\infty x^{n-1} \Sigma_{t,t+x} dx$ exist for $n = 1, 2$.

**Proposition 5**. *The first two moments $\bar{x}_t$ and $v_t$ of an admissible arbitrage-free term structure satisfy*

$$d\bar{x}_t = (r_t \bar{x}_t - 1)dt + \bar{\Sigma}_t(dW_t + \lambda_t dt), \tag{76}$$

*and*

$$dv_t = \left(r_t(v_t - \bar{x}_t^2) - \bar{\Sigma}_t^2\right) dt + 2\left(\bar{\Sigma}_t^{(1)} - \bar{x}_t \bar{\Sigma}_t\right)(dW_t + \lambda_t dt), \tag{77}$$

*where* $\bar{\Sigma}_t = \int_0^\infty \Sigma_{t,t+x} dx$ *and* $\bar{\Sigma}_t^{(1)} = \int_0^\infty x \Sigma_{t,t+x} dx$.

The proof of Proposition 5 can be found in Brody and Hughston (2001a). There is a critical value $\bar{x}_t^*$ for the first moment given by

$$\bar{x}_t^* = \frac{1}{r_t}(1 - \lambda_t \bar{\Sigma}_t), \tag{78}$$

such that when $\bar{x}_t > \bar{x}_t^*$ the drift of $\bar{x}_t$ is positive, and the drift increases as $\bar{x}$ increases. When $\bar{x}_t < \bar{x}_t^*$, the drift of $\bar{x}_t$ is negative, and the drift decreases further as $\bar{x}_t$ decreases. The process $v_t - \bar{x}_t^2$ measures the extent to which the distribution deviates from the 'flat' term structure. The second moment $v_t$ of the term structure has a positive drift providing

$$v_t - \bar{x}_t^2 > \frac{1}{r_t}\left(\bar{\Sigma}_t^2 - 2\lambda_t(\bar{\Sigma}_t^{(1)} - \bar{x}_t \bar{\Sigma}_t)\right). \tag{79}$$

The first moment $\bar{x}_t$ has the natural financial interpretation of being the value at time $t$ of a perpetual annuity paid on a continuous basis. In particular, a short calculation making use of an integration by parts shows that

$$\bar{x}_t = \int_0^\infty B_{tx} dx, \tag{80}$$

corresponding to an annuity of one unit of cash per year paid continuously in perpetuity.

Higher moments of the term structure density can then be interpreted in terms of the duration, convexity, etc., of the annuity—in other words, as a measure of the sensitivity of the value of the annuity to an overall change in interest rate levels. For example, if we write

$$B_{tx} = e^{-xr_t(x)}, \tag{81}$$

where $r_t(x)$ is the continuously compounded rate at time $t$ for tenor $x$, then under a small parallel shift $\Delta r$ in the yield curve given by

$$r_t(x) \longrightarrow r_t(x) + \Delta r, \tag{82}$$

we have, to first order,

$$B_{tx} \longrightarrow (1 - x\Delta r)B_{tx}. \tag{83}$$

Therefore, to first order the value of the annuity changes according to the scheme

$$\bar{x}_t \longrightarrow \bar{x}_t - \tfrac{1}{2}\Delta r \int_0^\infty x^2 \rho_t(x)dx, \tag{84}$$

where in obtaining the second term we use an integration by parts. It follows then that:

**Proposition 6**. *Under a parallel shift in the yield curve the change $\Delta\bar{x}_t$ in the value of the perpetual is*

$$\Delta\bar{x}_t = -D_t\bar{x}_t\Delta r, \tag{85}$$

*where the duration $D_t$ of the perpetual annuity is given by*

$$D_t = \frac{\tfrac{1}{2}\int_0^\infty x^2 \rho_t(x)dx}{\int_0^\infty x\rho_t(x)dx}. \tag{86}$$

## 6. The information content of the term structure

Now we introduce another important example of a functional of the term structure, the Shannon entropy of the density function $\rho_t(x)$. This is defined by

$$S_t[\rho] = -\int_0^\infty \rho_t(x) \ln \rho_t(x)dx. \tag{87}$$

There are other measures of entropy that may also have applications in interest rate theory, though the Shannon entropy is distinguished amongst these by its simplicity. Because $\rho_t(x)$ has dimensions of inverse time, $S_t[\rho]$ is defined only up to an overall additive constant. The difference of the entropies associated with two yield curves therefore has an invariant significance. One can think of $S_t[\rho]$ as being a measure of the 'information content' of the term structure at time $t$. In particular, the higher the value of $S_t[\rho]$, the lower the information content (see, e.g., Jaynes 1982, 1983, Cover & Thomas 1991).

Since $\rho_t(x)$ is subject to a dynamical law, we can infer a corresponding dynamics for the entropy, given as follows.

**Proposition 7**. *The entropy associated with an admissible arbitrage-free term structure dynamics obeys the evolutionary law*

$$dS_t = \left(r_t(S_t + \ln r_t - 1) + \tfrac{1}{2}\Gamma_t\right)dt + \left(\int_0^\infty \nu_t(x)s_t(x)dx - \bar{\nu}_t S_t\right)dW_t^* \tag{88}$$

*where $dW^* = dW_t + \lambda_t dt$, $s_t(x) = -\rho_t(x)\ln\rho_t(x)$ is the entropy density, and the process $\Gamma_t$ is defined by*

$$\Gamma_t = \int_0^\infty (\nu_t(x) - \bar{\nu}_t)^2 \rho_t(x)dx. \tag{89}$$

*Proof.* By Ito's lemma we have

$$d(\rho_t(x) \ln \rho_t(x)) = (1 + \ln \rho_t(x))d\rho_t(x) + \frac{1}{\rho_t(x)}(d\rho_t(x))^2. \tag{90}$$

It follows by use of (47) that $s_t(x)$ satisfies

$$\begin{aligned} ds_t(x) = & -\left((1 + \ln \rho_t(x))\left(r_t \rho_t(x) + \partial_x \rho_t(x)\right) + \rho_t(x)(\nu_t(x) - \bar{\nu}_t)^2\right) dt \\ & - \rho_t(x)(1 + \ln \rho_t(x))(\nu_t(x) - \bar{\nu}_t)(dW_t - \bar{\nu}_t dt). \end{aligned} \tag{91}$$

The desired result is obtained if we integrate over $x$ and make use of the integral identities

$$\int_0^\infty \partial_x \rho_t(x) dx = -r_t \tag{92}$$

and

$$\int_0^\infty \ln \rho_t(x) \partial_x \rho_t(x) dx = r_t(1 - \ln r_t). \tag{93}$$

$$\diamond$$

The principle of entropy maximisation can be used quite effectively as the basis for a new yield curve calibration methodology. In particular, given a set of data points on a yield curve, the least biased and hence most plausible term structure can be determined by maximising the Shannon entropy subject to the given data constraints.

The general idea behind the maximisation of entropy under constraints can be sketched as follows. Suppose that, given a function $H(x)$ of a random variable $x$, we are told that the expectation of $H(x)$ with respect to an unknown distribution with density $\rho(x)$ is $U$, i.e.,

$$\int_0^\infty H(x)\rho(x) dx = U. \tag{94}$$

The aim then is to find the density $\rho(x)$ that is least biased and yet consistent with the information (94). In other words, we wish to eliminate any superfluous information in $\rho(x)$. In addition to (94), we also have the normalisation condition

$$\int_0^\infty \rho(x) dx = 1. \tag{95}$$

Subject to the constraints (94) and (95) we then determine the density $\rho(x)$ that maximises the entropy. This is most readily carried out by introducing Lagrange multipliers $\lambda$ and $\nu$, and considering the variational relation

$$\frac{\delta}{\delta \rho}\left(-\rho \ln \rho - \lambda \rho H - \nu \rho\right) = 0. \tag{96}$$

The solution is

$$\rho(x) = \exp\left(-\lambda H(x) - \nu - 1\right), \tag{97}$$

where $\lambda$ and $\nu$ are determined implicitly by (94) and (95).

Let us illustrate the idea by considering the situation in which we are given a set of data points on the yield curve together with the value of a perpetual annuity. The problem is to calibrate the initial term structure to the given data.

This example is interesting because if we are given only the value $\bar{x}_0$ of the perpetual annuity, then the maximum entropy term structure is

$$\rho_0(x) = Re^{-Rx}, \tag{98}$$

where $R = 1/\bar{x}_0$, and thus $P_{0x} = e^{-Rx}$ for the discount function. Therefore, we see that it is the annuity constraint that leads to the desired exponential 'die-off' of the discount function. This feature is preserved in the more elaborate examples we discuss below, where bond data-points are introduced as well.

In the more general situation, the information of the data can be captured by saying that the bond prices with a given set of tenors $x_i$ $(i = 1, 2, \cdots, r)$ are observed to be $B_{0x_i} = \eta_i$. In addition, we have the initial value $\bar{x}_0 = \xi$ of the perpetual annuity. Subject to these constraints, it is a straightforward exercise to verify that the maximum entropy term structure is determined by the variational principle

$$\frac{\delta}{\delta\rho}\left(-\rho(x)\ln\rho(x) - \lambda\rho(x)x - \sum_{i=1}^{r}\mu^i\rho(x)I_{x_i}(x) - \nu\rho(x)\right) = 0, \tag{99}$$

where the step function $I_{x_i}(x)$ is given by (43), so $I_{x_i}(x) = 1$ for $x \geq x_i$ and vanishes otherwise. The parameters $\lambda$, $\mu^i$ and $\nu$ are Lagrange multipliers to be determined by the normalisation condition (33) and data constraints

$$\int_{x=0}^{\infty} x\rho(x)dx = \xi, \quad \text{and} \quad \int_{x=0}^{\infty} I_{x_i}(x)\rho(x)dx = \eta_i. \tag{100}$$

The solution of the variational problem (99) is

$$\rho(x) = \frac{1}{Z(\lambda, \mu)}\exp\left(-\lambda x - \sum_{i=1}^{r}\mu^i I_{x_i}(x)\right), \tag{101}$$

where the normalisation factor $Z(\lambda, \mu) = e^{-1-\nu}$ is determined by the integral

$$Z(\lambda, \mu) = \int_0^{\infty}\exp\left(-\lambda x - \sum_{i=1}^{r}\mu^i I_{x_i}(x)\right)dx. \tag{102}$$

The Lagrange multipliers are then determined implicitly by the following relations:

$$-\frac{\partial \ln Z}{\partial \lambda} = \xi \quad \text{and} \quad -\frac{\partial \ln Z}{\partial \mu^i} = \eta_i. \tag{103}$$

As a consequence of (101) we see that pointwise calibration to the discount bond prices, along with the information of the price of the annuity, gives a piecewise exponential term structure density function. Clearly, if there is further information at our disposal, then that can also be included in the system of constraints so that all available information is used efficiently in the calibration procedure. For example, if a set of data points as well as the second moment of the term structure density is known, then the resulting term structure becomes piecewise Gaussian.

We now consider in more detail the case where the observed data consist of just two pieces of information—namely, the bond price $P_{0T_1}$ for a fixed maturity date $T_1$, and the value $\xi = \bar{x}_0$ of the perpetual annuity. This is, of course, a rather artificial

example; nevertheless it serves to illuminate some of the main points of the procedure and has the advantage of being analytically tractable.

It follows from the proceeding discussion that the variational problem in this example implies the existence of three rates $r_0$, $r_1$, and $R$ such that the term structure density is

$$\rho(x) = \begin{cases} r_0 e^{-Rx} & \text{for} \quad 0 \le x < T_1 \\ r_1 e^{-Rx} & \text{for} \quad T_1 \le x < \infty. \end{cases} \tag{104}$$

The relevant constraints are given by:

$$\int_0^{T_1} \rho(x)dx = 1 - P_{0T_1} \tag{105}$$

for the bond price;

$$\int_0^\infty \rho(x)dx = 1 \tag{106}$$

for the normalisation; and

$$\int_0^\infty x\rho(x)dx = \xi \tag{107}$$

for the perpetual annuity. A short calculation making use of (104) shows that these relations reduce to:

$$1 - \frac{r_0}{R}\left(1 - e^{-RT_1}\right) = P_{0T_1}, \tag{108}$$

$$\frac{r_1 - r_0}{R}e^{-RT_1} + \frac{r_0}{R} = 1, \tag{109}$$

$$\frac{r_1 - r_0}{R}e^{-RT_1}\left(T_1 + \frac{1}{R}\right) + \frac{r_0}{R^2} = \xi. \tag{110}$$

Clearly, given $P_{0T_1}$ and $\xi$, we can proceed, by use of (108), (109) and (110), to infer values of $r_0$, $r_1$, and $R$. In particular, equation (108) allows us to deduce the bond price $P_{0T_1}$ if we are given $r_0$ and $R$, whereas we can use (109) to eliminate $r_1$ in (110) to obtain

$$\frac{1}{R} + T_1\left(1 - \frac{r_0}{R}\right) = \xi \tag{111}$$

for the value of the perpetual in terms of $r_0$ and $R$. Alternatively, given the initial short rate $r_0$ and the value of the perpetual $\xi$ we have

$$R = \frac{1 - r_0 T_1}{\xi - T_1}. \tag{112}$$

This value of $R$ can then be inserted in (108) to determine the bond price. The scale factor $r_1$ is given by

$$r_1 = RP_{0T_1}e^{RT_1}, \tag{113}$$

and thus we obtain

$$\rho(x) = \begin{cases} r_0 e^{-Rx} & (0 \le x < T_1) \\ RP_{0T_1}e^{-R(x-T_1)} & (T_1 \le x < \infty), \end{cases} \tag{114}$$

for the term structure density, and

$$P_{0x} = \begin{cases} 1 - \frac{r_0}{R}\left(1 - \mathrm{e}^{-Rx}\right) & (0 \le x < T_1) \\ P_{0T_1}\mathrm{e}^{-R(x-T_1)} & (T_1 \le x < \infty), \end{cases} \tag{115}$$

for the discount function, from which yield curve $R_{0x}$ can be constructed via the standard prescription

$$R_{0x} = -\frac{\ln P_{0x}}{x}, \tag{116}$$

and it should be evident by inspection that $R_{0x}$ is continuous in $x$.

In this example we can alternatively regard the short rate $r_0$ and the bond price $P_{0T_1}$ as the actual 'independent' data. Then (108) can be used to deduce $R$, which allows us to infer the annuity price $\xi$ by use of (111). This illustrates the point that, although we assume from the outset the *existence* of a perpetual, we can infer an implied value of that instrument by the use of other market data (e.g., the short rate).

Essentially the same idea carries forward in a consistent way in the case where we have multiple data points for the bond prices, for a given set of $n$ maturity dates $T_j$ $(j = 1, 2, \cdots, n)$, and we are led to a simple iterative algorithm for determining the term structure in terms of the short rate and the specified bond data points.

**Proposition 8**. *Given a set of bond prices $P_{0T_j}$ $(j = 1, 2, \cdots, n)$ and the existence of the value of the perpetual annuity, the maximum entropy term structure density function is*

$$\rho(x) = \sum_{k=0}^{n} I_{T_k T_{k+1}}(x) r_k \mathrm{e}^{-Rx}. \tag{117}$$

*Here $T_0 = 0$, $T_{n+1} = \infty$, $I_{T_k T_{k+1}}(x) = 1$ if $x \in [T_k, T_{k+1})$ and vanishes otherwise, $r_0$ is the short rate, and*

$$r_k = R\frac{P_{0T_k} - P_{0T_{k+1}}}{\mathrm{e}^{-RT_k} - \mathrm{e}^{-RT_{k+1}}}. \tag{118}$$

*The value of $R$ is determined from equation (108). The corresponding discount function $P_{0x}$ is given by*

$$P_{0x} = P_{0T_k} - \frac{r_k}{R}(\mathrm{e}^{-RT_k} - \mathrm{e}^{-Rx}) \tag{119}$$

*for $x \in [T_k, T_{k+1})$.*

*Proof.* To see this, we insert the piecewise exponential density function (117) into a series of constraints of the form (105) for the bond prices, together with the normalisation constraint (106) and the perpetual constraint (107). Then the bond price constraints give rise to a set of relations of the form

$$\frac{r_k}{R}(\mathrm{e}^{-RT_k} - \mathrm{e}^{-RT_{k+1}}) = P_{0T_k} - P_{0T_{k+1}}, \tag{120}$$

for $k = 0, 1, \cdots, n - 1$. In particular, for $k = 0$, we recover (108), which can be used to solve for $R$ in terms of the short rate $r_0$ and the bond price $P_{0T_1}$. Then, by substitution of this in (120) for general $k$, and the use of further bond price data, we obtain the other

rates $r_k$ $(k \neq n)$. As for $r_n$, we note that if we divide the integration range in (105) into two regions $[0, T_n]$ and $[T_n, \infty]$, then the normalisation condition becomes

$$\frac{r_n}{R} \mathrm{e}^{-RT_n} = P_{0T_n}, \tag{121}$$

which determines $r_n$ in terms of $R$ and $P_{0T_n}$. Finally, substitution of these results in the perpetual constraint

$$\frac{1}{R} \sum_{k=1}^{n} (r_k - r_{k-1}) \mathrm{e}^{-RT_k} \left( T_k + \frac{1}{R} \right) + \frac{r_0}{R^2} = \xi \tag{122}$$

allows the implied value $\xi$ of the perpetual annuity to be determined from the short rate $r_0$ and the bond price data $P_{0T_j}$. The discount function can be determined by use of the fact that

$$
\begin{aligned}
1 - P_{0x} &= \int_0^x \rho(u) du \\
&= \int_0^{T_k} \rho(u) du + r_k \int_{T_k}^x \mathrm{e}^{-Ru} du \\
&= 1 - P_{0T_k} + r_k \int_{T_k}^x \mathrm{e}^{-Ru} du,
\end{aligned}
\tag{123}
$$

when $x \in [T_k, T_{k+1})$. ◇

## 7. Canonical term structures

As an interesting example of a class of models that arises as a consequence of the maximisation of an entropy functional under constraints, we let the term structure density be of the form

$$\rho_t(T - t) = \frac{\exp\left(-g_{tT} - \theta_t h_{tT}\right)}{\int_{u=t}^{\infty} \exp\left(-g_{tu} - \theta_t h_{tu}\right) du}, \tag{124}$$

where $\theta_t$ is a one-dimensional Ito process, and the functions $g_{tT}$ and $h_{tT}$ are deterministic, defined over the range $0 \leq t \leq T < \infty$. At each time $t$ the term structure density thus defined belongs to an exponential family parameterised by the value of $\theta_t$. If we set

$$Z(\theta) = \int_{u=t}^{\infty} \exp\left(-g_{tu} - \theta_t h_{tu}\right) du, \tag{125}$$

then we find that all the moments of the function $h_{tT}$ can be determined from the generating function $Z(\theta)$ by formal differentiation. For example, for the first moment of $h_{tT}$ we have

$$\int_{u=t}^{\infty} h_{tu} \rho_t(u - t) du = -\frac{\partial \ln Z(\theta)}{\partial \theta}. \tag{126}$$

The corresponding bond price system can then be written in the Flesaker-Hughston form

$$P_{tT} = \frac{\int_{u=T}^{\infty} N_{tu} du}{\int_{u=t}^{\infty} N_{tu} du}, \tag{127}$$

where $N_{tT} = \exp(-g_{tT} - \theta_t h_{tT})$. By Ito's lemma, it follows that $N_{tT}$ satisfies

$$\frac{dN_{tT}}{N_{tT}} = -\left(\dot{g}_{tT} + \theta_t \dot{h}_{tT}\right) dt - h_{tT} d\theta_t + \tfrac{1}{2} h_{tT}^2 (d\theta_t)^2, \tag{128}$$

where the dot indicates partial differentiation with respect to $t$, so $\dot{g}_{tT} = \partial_t g_{tT}$ and $\dot{h}_{tT} = \partial_t h_{tT}$. We assume that the trajectory $\theta_t$ of the canonical parameter satisfies a stochastic equation of the form

$$d\theta_t = \alpha_t dt + \beta_t dW_t. \tag{129}$$

The no-arbitrage condition implies that $N_{tT}$ is a positive martingale. Therefore, the drift of $N_{tT}$ vanishes for all $T$:

$$\dot{g}_{tT} + \dot{h}_{tT}\theta_t + \alpha_t h_{tT} = \tfrac{1}{2}\beta_t^2 h_{tT}^2. \tag{130}$$

This relation implies that the processes $\alpha_t$ and $\beta_t$ determining the dynamics of $\theta_t$ are of the form

$$\alpha_t = A_t \theta_t + B_t, \quad \text{and} \quad \tfrac{1}{2}\beta_t^2 = C_t \theta_t + D_t \tag{131}$$

where the functions $A_t$, $B_t$, $C_t$ and $D_t$ are deterministic. It follows that $\theta_t$ is a square-root process. Substitution of these equations into (130) gives a set of Bernoulli equations of the form

$$\dot{h}_{tT} + A_t h_{tT} - C_t h_{tT}^2 = 0 \tag{132}$$

for $h_{tT}$ and

$$\dot{g}_{tT} + B_t h_{tT} - D_t h_{tT}^2 = 0 \tag{133}$$

for $g_{tT}$. The general solution of (132) is

$$\frac{1}{h_{tT}} = -\exp\left(\int_0^t A_u du\right)\left(\int_0^t \frac{C_u}{\exp(\int_0^u A_v dv)} du + E_T\right), \tag{134}$$

where $E_T$ is an function of $T$, determined by the initial term structure.

To proceed further, let us consider the special case where $D_t = 0$ and $\theta_t$ is positive and mean-reverting. Then $B_t$ and $C_t$ are both positive and $A_t$ is negative, and for $g_{tT}$ we have

$$g_{tT} = -\int_0^t B_u h_{uT} du + F_T, \tag{135}$$

where $F_T$ is another arbitrary function. In the elementary case where $A_t$, $B_t$ and $C_t$ are constants, the functions $h_{tT}$ and $g_{tT}$ are given by

$$h_{tT} = \frac{A}{C - G_T e^{At}} \tag{136}$$

and

$$g_{tT} = \frac{B}{C} \ln\left(\frac{G_T - Ce^{-At}}{G_T - C}\right) + F_T, \tag{137}$$

where $G_T = AE_T + C$. The condition that $h_{tT}$ should be positive ensures that $G_T$ is of the form $G_T = CH_T \mathrm{e}^{-AT}$ where the function $H_T$ satisfies $H_T > 1$ but is otherwise arbitrary. For $N_{tT}$ we then obtain:

$$N_{tT} = \left( \frac{H_T - \mathrm{e}^{AT}}{H_T - \mathrm{e}^{A(T-t)}} \right)^{\frac{B}{C}} \exp \left( \frac{A\theta_t}{C(H_T \mathrm{e}^{-A(T-t)} - 1)} - F_T \right). \qquad (138)$$

The function $F_T$ is then determined by the specification of the initial term structure. In particular, because $N_{0T} = \rho_0(T)$, we obtain

$$N_{tT} = \rho_0(T) \left( \frac{H_T - \mathrm{e}^{AT}}{H_T - \mathrm{e}^{A(T-t)}} \right)^{\frac{B}{C}} \exp \left( \frac{A\theta_t}{C(H_T \mathrm{e}^{-A(T-t)} - 1)} - \frac{A\theta_0}{C(H_T \mathrm{e}^{-AT} - 1)} \right). \qquad (139)$$

## Acknowledgments

## References

[1] Bhattacharyya, A. 1943 On a measure of divergence between two statistical populations defined by their probability distributions *Bull. Calcutta Math. Soc.* **35**, 99–109.

[2] Björk, T. & Christensen, B. J. 1999 Interest rate dynamics and consistent forward rate curves. *Mathematical Finance* **9**, 323–348.

[3] Björk, T. & Gombani, A. 1999 Minimal realizations of interest rate models. *Finance and Stochastics* **3**, 413–432.

[4] Björk, T. 2001 A geometric view of interest rate theory. In *Option pricing, interest rates and risk management, Handb. Math. Finance*, Cambridge: Cambridge University Press.

[5] Brody, D. C. 2000 *Modern Mathematical Theory of Finance*, Tokyo: Nippon-Hyoronsya.

[6] Brody, D. C. & Hughston, L. P. 2001a Interest rates and information geometry. *Proc. Roy. Soc. London* A**457**, 1343–1364.

[7] Brody, D. C. & Hughston, L. P. 2001b Applications of information geometry to interest rate theory, In *Disordered and Complex Systems*, P Sollich, ACC Coolen, LP Hughston, RF Streater (eds), New York: AIP Publishing.

[8] Cover, T. M. & Thomas, J. A. 1991 *Elements of Information Theory*, New York: John Wiley & Sons.

[9] Filipović, D. 2001 *Consistency Problems for Heath-Jarrow-Morton Interest Rate Models*, Lecture Notes in Mathematics **1760**, Berlin: Springer-Verlag.

[10] Flesaker, B. & Hughston, L. P. 1996 Positive Interest *Risk Magazine* **9**, 46–49; reprinted in *Vasicek and Beyond*, L.P. Hughston (ed), London: Risk Publications (1996).

[11] Flesaker, B. & Hughston, L. P. 1997 International models for interest rates and foreign exchange *Net Exposure* **3**, 55–79; reprinted in *The New Interest Rate Models*, L.P. Hughston (ed), London: Risk Publications (2000).

[12] Flesaker, B. & Hughston, L. P. 1998 Positive Interest: An Afterword, in *Hedging with Trees*, Broadie, M. & Glasserman, P. (eds), London: Risk Publications.

[13] Heath, D., Jarrow, R. & Morton, A. 1992 Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* **60**, 77–105.

[14] Hunt, P. J. & Kennedy, J. E. 2000 *Financial Derivatives in Theory and Practice*, Chichester: John Wiley & Sons.

[15] Jaynes, E. T. 1982 On the rationale of maximum entropy methods. *Proc. IEEE* **70**, 939–952.

[16] Jaynes, E. T. 1983 *Papers on probability, statistics and statistical physics: edited and with an introduction by R. D. Rosenkrantz*, Dordrecht: D. Reidel Publishing Co.

[17] Musiela, M. & Rutkowski, M. 1997 *Martingale Methods in Financial Modelling*, Berlin: Springer-Verlag.

[18] Rogers, L. C. G. 1997 The potential approach to the term structure of interest rates and foreign exchange rates *Math. Finance* **7**, 157–176; reprinted in *The New Interest Rate Models*, L.P. Hughston (ed), London: Risk Publications (2000).

[19] Rutkowski, M. 1997 A note on the Flesaker-Hughston model of the term structure of interest rates *Applied Math. Finance* **4**, 151–163; reprinted in *The New Interest Rate Models*, L.P. Hughston (ed), London: Risk Publications (2000).