Financial Options Research Centre

University of Warwick

Valuing Interest Rate Options via a Primitive Theory of the Term Structure

Andrew P Carverhill

August 1989

FORC Preprint: 89/5

VALUING INTEREST RATE OPTIONS VIA A PRIMITIVE THEORY OF THE TERM STRUCTURE

1. INTRODUCTION

The Primitive Theory of the term structure of interest rates has been presented in the paper [C 1989] It is so called because it is based on a primitive, intuitively easy assumption about how the term structure, or yield curve, evolves. In fact in [C 1989] the theory is worked out in 1-factor form, that factor being the short interest rate, and in this form of the theory, the primitive assumption is as follows: the yield on a bond at a short time into the future will be given by the forward yield implied by current prices, plus a random element which is perfectly correlated to the change in the short rate. This assumption will be stated in full mathematical brutality in Section 2 below.

The Primitive Theory is closely related to the theories of [HL] and [HJM], which also deals with the evolution of the term structure. Also the Primitive Theory uses the ideas of the factor analysis of the evolution, which are presented in [S 1989b] and [D]. "Factor Analysis" is a technique which naively identifies the principal components of the evolution. By "naively" here we mean that the technique does not rely on any economic theories about the term structure, but merely looks at the magnitudes involved and calculates various covariances and eigenvectors. In fact one might say that the Primitive Theory is a reformulation of the theory of [HJM], using the frame-of-reference of Factor Analysis, which makes it easy to understand in terms of elasticities, covariances, etc.

A plan of this paper is as follows:

In Section 2 we present the Primitive Theory in 1-factor short-rate form as described above. Also we present a number of reformulations of this theory, which are completely equivalent to each other, but which are different in their convenience of use in various situations, i.e. different formulations are appropriate for theoretical study, empirical estimation, option evaluation etc. We admit that our constant changing between these formulations makes the Primitive Theory perhaps look esoteric.

In Section 3 we present the 2-factor long-short-rate form of the theory. This form of the theory takes the long and short rates as independent driving processes for the term structure, and shows how to obtain intermediate rates, rather as is done in [SS], except that the Primitive Theory also uses the current term structure in its prediction. We also relate the Primitive Theory to the Factor Analysis of the evolution of the term structure. This Factor Analysis can in principle provide 3-factor and higher-factor forms of the theory, and the driving processes do not have to be directly identifiable in economic terms.

In Section 4 we show how to value interest rate contingent claims based on the Primitive Theory, concentrating on the 1-factor short-rate form of the theory. Also we emphasise interest rate caps and floors in this section. These instruments provide an appropriate test of the short-rate form of the theory because they involve short rates, and they live for a sufficiently long time (up to 5 years) that the current term structure, which is taken into account by the theory, is an important determinant of their value.

2. THE 1-FACTOR SHORT-RATE PRIMITIVE THOERY OF THE TERM STRUCTURE

We represent the term structure of interest rates at time t by the set of prices $\{P(t,q) : q \geq t\}$ of pure discount bonds. The pure discount bond with price P(t,q) at time t, pays 1 at time q, and there is no risk of default on this payment. Of course these bonds do not exist for all maturities in the market, but their virtual prices can be estimated from the prices of Government bonds, using the procedure described in the paper [S 1989a]. That paper also shows that the British Gilt-Edged market behaves efficiently, as though these bonds do exist.

We can also represent the term structure at time t as the collection of spot rates or yields $\{p(t,q) : q \geq t\}$, where

$$p(t,q) = -\frac{\log P(t,q)}{(q-t)}. \tag{2.1}$$

Other magnitudes that we will use in this paper are the forward price G(t,s,q) at time t of the discount bond to be purchased at time s and which matures at time q, and the corresponding forward rate g(t,s,q). These are given by the formulae

$$G(t,s,q) = P(t,q)/P(t,s), \tag{2.2}$$

$$g(t,s,q) = -\frac{\log G(t,s,q)}{(q-s)} \tag{2.3}$$

(n.b. The forward rate g is different from the "instantaneous" forward rate f, which we introduce later.)

As we have already said, in this section we deal with the 1-factor form of the Primitive Theory, in which the driving process is taken to be the short rate $r_t \equiv p(t,t)$. We assume that r_t itself is driven by the autonomous stochastic equation

$$dr_t = \xi(r_t)dt + \eta(r_t)dB_t, \tag{2.4}$$

in which B_t is a standard Brownian Motion, and the coefficients ξ and η must be determined empirically, but are allowed to depend on the short rate. (n.b. The notation ξ and η is unusual, but is chosen to prevent a clash later.)

We will use equation (2.4) in the "fattened up" version given by equation (2.5) below, and this version gives the key to understanding the stochastic equation. Equation (2.5) holds for small ε, and is

$$\Delta_t^{t+\varepsilon} r = \xi(r_t)\Delta_t^{t+\varepsilon} t + \eta(r_t)\Delta_t^{t+\varepsilon} B, \tag{2.5}$$

where

$$\Delta_t^{t+\varepsilon} r = r_{t+\varepsilon} - r_t,$$

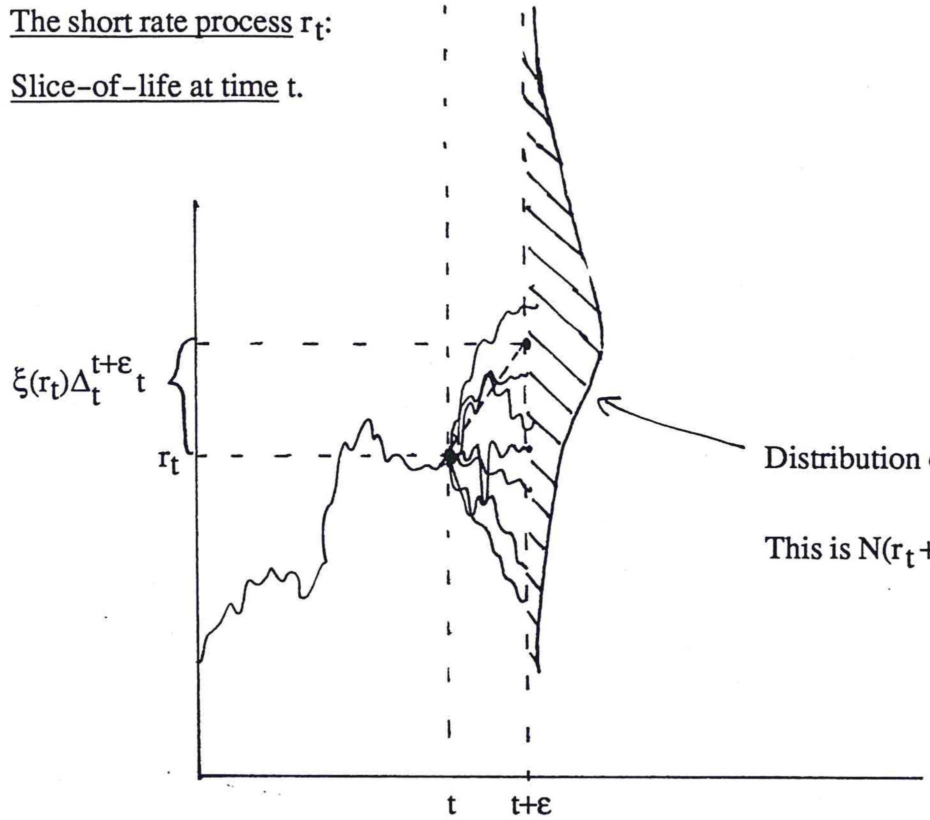
$$\Delta_t^{t+\varepsilon} t = (t+\varepsilon) - t \equiv \varepsilon,$$

$$\Delta_t^{t+\varepsilon} B = B_{t+\varepsilon} - B_t.$$

(Note that the increment $B_{t+\varepsilon} - B_t$ has distribution $N(0,\sqrt{\varepsilon})$, i.e. it is normally distributed with mean 0 and standard diviation $\sqrt{\varepsilon}$.) In Equation (2.5) the differentials of Equation (2.4) have been replaced by increments, or "differences", over the short time interval [t, t+ε]. Equation (2.5) tells us that over this interval the increment in the short rate is composed of a drift at rate $\xi(r_t)$ and a noise with volatility $\eta(r_t)$, so that this increment is normally distributed with mean $\xi(r_t)\Delta_t^{t+\varepsilon} t \equiv \xi(r_t)\varepsilon$, and standard deviation $\eta(r_t)$ {standard derivation of $\Delta_t^{t+\varepsilon} B\} \equiv \eta(r_t)\sqrt{\varepsilon}$. Technical purists will note that (2.5) can only hold approximately given (2.4), with error which vanishes as we go from the "difference" to the "differential", i.e. as we let ε tend to zero. However, we can and will ignore this error in this and other "fattened up" equations.

<u>The short rate process</u> r_t:

<u>Slice-of-life at time</u> t.



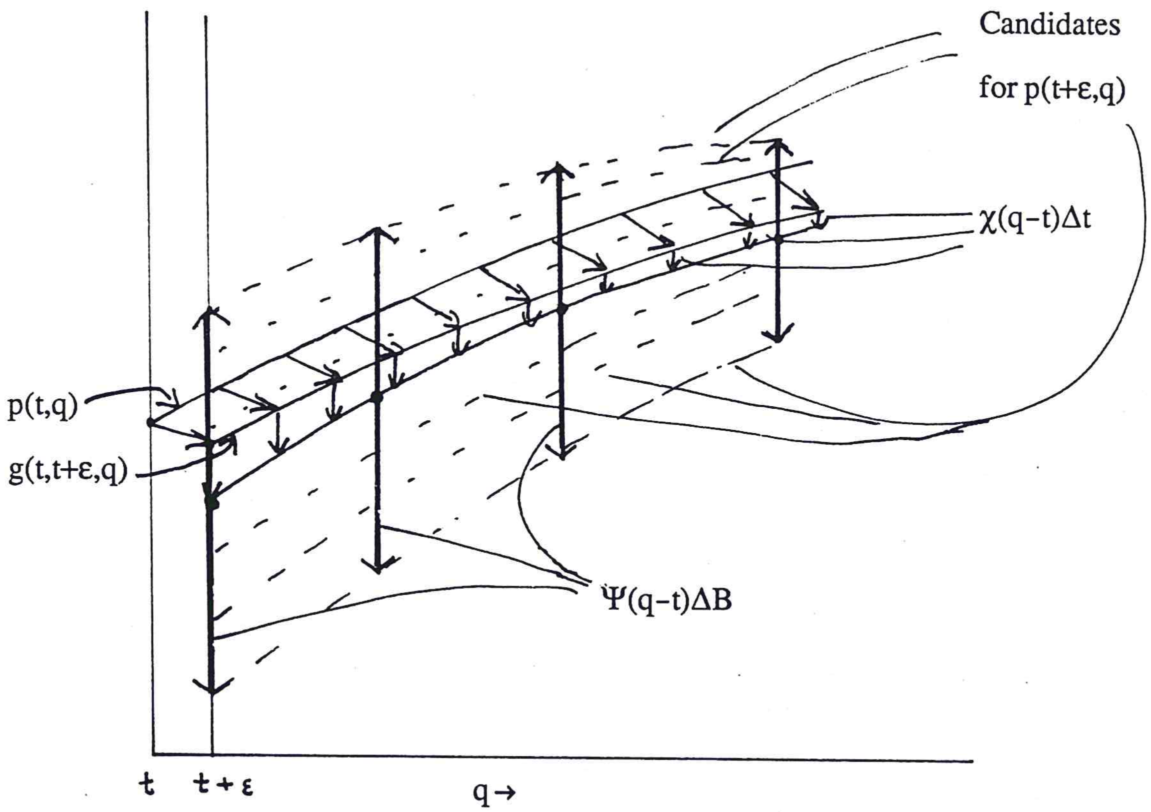
Distribution of $r_{t+\varepsilon}$ given r_t.

This is $N(r_t + \xi(r_t)\varepsilon, \eta(r_t)\sqrt{\varepsilon})$

The basic assumption of the short-rate version of the Primitive Theory is as follows: Assume we are currently at time t. Then at a short time t+ε in the future the spot rate $p(t+\varepsilon, q)$ on the time q maturity bond will be given by the forward rate $g(t, t+\varepsilon, q)$, plus a random element which is perfectly correlated to the short rate. This assumption is formulated below in Equation (2.6), in which the random element is given in terms of a drift χ and noise Ψ, which are allowed to depend on the time to maturity (q−t) and the current short rate r_t. The equation is

$$p(t+\varepsilon,q) = g(t,t+\varepsilon,q) + \chi_{r_t}(q-t)\Delta_t^{t+\varepsilon} t + \Psi_{r_t}(q-t)\Delta_t^{t+\varepsilon} B. \tag{2.6}$$

The evolution described by Equation (2.6) can be represented by the following diagram:



When working with the Primitive Theory we must cast it into a number of reformulations, which we describe below. These reformulations tend to make the theory look ever more complicated and "esoteric". However, we hope the reader will agree that the basic formulation given by Equations (2.5) and (2.6) is "primitive" in the sense of being intuitively easy. We mention two additional conditions which we will impose in the Primitive Theory, and which are justified in [C 1989]. One is a consistency condition between Equations (2.5) and (2.6) and is given by

$$\xi(r) = \chi_r(0), \quad \eta(r) = \Psi_r(0). \tag{2.7}$$

The other condition prevents arbitrage between bonds of different maturities and is that

$$\chi_r(s) - \tfrac{1}{2}s\Psi_r(s)^2 = \gamma_r\Psi_r(s), \tag{2.8}$$

where γ_r can depend on the short rate r, but is independent of s. This γ_r can be interpreted as the risk premium in the bond market, as will become clear below. It is the rate of return on the bond, in excess of the riskless rate of return, i.e. the short rate, which is required to conpensate for each unit of the bonds volatility.

Next we give our reformulations of the short-rate version of the Primitive Theory. These are all justified in [C 1989]. It turns out that each formulation is suited to a particular aspect of the Primitive Theory, and each one can be regarded as "primitive" from a certain perspective. The "esoteric" nature of the theory arises from our frequent switching between these formulations.

Our first reformulation is to substitute for the random element $\Delta_t^{t+\varepsilon}B$ in (2.6) using (2.5) to obtain

$$p(t+\varepsilon,q) = g(t,t+\varepsilon,q) + \tilde{\chi}_{r_t}(q-t)\Delta_t^{t+\varepsilon}t + \tilde{\Psi}_{r_t}(q-t)\Delta_t^{t+\varepsilon}r \qquad (2.9)$$

where

$$\left.\begin{aligned} \tilde{\chi}_r(q-t) &= \chi_r(q,t) - (\xi(r)/_{\eta(r)})\Psi_r(q-t) \\ \tilde{\Psi}_r(q-t) &= (1/_{\eta(r)})\Psi_r(q-t). \end{aligned}\right\} \qquad (2.10)$$

This formulation has a number of advantages over the original one, and these make it suitable as a starting point for empirical estimation of the Primitive Theory. First, (2.9) makes explicit the fact that the random element in the evolution of the maturity-time q interest rate, is the evolution of the short interest rate. In fact, if we put $\Delta_t^{t+\varepsilon}p(q)$ = $p(t+\varepsilon,q) - g(t,t+\varepsilon,q)$, then $\tilde{\chi}$ and $\tilde{\Psi}$ can be interpreted in terms of the correlation between $\Delta_t^{t+\varepsilon}p$ and $\Delta_t^{t+\varepsilon}r$, or the elasticity between the q maturity rate and the short rate. Also $\tilde{\chi}$ and $\tilde{\Psi}$ can be estimated empirically by regression; all the other terms in (2.9) can be calculated from estimates of the term structure, and for each time-to-maturity q-t, the coefficients $\tilde{\chi}(q-t)$ and $\tilde{\Psi}(q-t)$ can be estimated using data for the other terms pertaining to a series starting times t. Having obtained $\tilde{\chi}$ and $\tilde{\Psi}$, then χ and Ψ can be calculated using the inverse of (2.10), which is

$$\left.\begin{aligned} \chi_r(q-t) &= \tilde{\chi}_r(q-t)+\xi(r), \\ \Psi_r(q-t) &= \eta(r)\tilde{\Psi}_r(q-t). \end{aligned}\right\} \tag{2.11}$$

In this reformulation, the condition (2.7) translates over to the condition

$$\tilde{\chi}_r(0) = 0, \quad \tilde{\Psi}_r(0) = 1, \tag{2.12}$$

and in fact (2.12) is more natural, and is clear from the easily established fact that

$$\Delta_t^{t+\varepsilon} p(q) \equiv p(t+\varepsilon,q) - g(t,t+\varepsilon,q) = \Delta_t^{t+\varepsilon} r + \{\text{small error}\}. \tag{2.13}$$

Note that in view of (2.7) it is not reasonable to assume that the coefficients χ and Ψ are independent of the short rate r, but it might not be unreasonable to assume that the coefficients $\tilde{\chi}$ and $\tilde{\Psi}$ are independent of r, and we implicitly assumed this in the previous paragraph. If χ and Ψ were independent of r, then the coefficients ξ and η would be also by (2.7), and then the short rate would not be mean-reverting.

Another reformulation in [C 1989] casts Equation (2.6) in terms of prices rather than rates. The reformulated Equation (2.6) can be any of the following:

$$P(t+\varepsilon,q) = G(t,t+\varepsilon,q)\,[1 + \nu_{r_t}(q-t)\Delta_t^{t+\varepsilon} t + \mu_{r_t}(q-t)\Delta_t^{t+\varepsilon} B], \tag{2.14}$$

$$P(t+\varepsilon,q)/P(t,q) = 1 + [\nu_{r_t}(q-t) + r_t]\Delta_t^{t+\varepsilon} t + \mu_{r_t}(q-t)\Delta_t^{t+\varepsilon} B, \tag{2.15}$$

$$dP(t,q)/P(t,q) = [\nu_{r_t}(q-t) + r_t]dt + \mu_{r_t}(q-t)dB_t, \tag{2.16}$$

where

$$\left.\begin{aligned}\nu_r(s) &= -s\chi_r(s) + \tfrac{1}{2}s^2\Psi_r(s)^2,\\ \mu_r(s) &= -s\Psi_r(s).\end{aligned}\right\} \qquad (2.17)$$

Of these Equations (2.14), (2.15), (2.16), the last is the most user-friendly. Note that (2.15) is just a "fattened up" version of (2.16). Also, to go from (2.14) to (2.15) we just expand the term $G(t,t+\varepsilon,q)$ with respect to ε, as explained in [C 1989]. Note also that it would perhaps be more usual to absorb the r_t term into the ν term in (2.16), but we prefer not to do this.

We will start from this formulation when we discuss option evaluation in Section 4. Also we can immediately give an advantage of this formulation, namely a rephrasing of condition (2.8), which relates to arbitrage-across maturities. This condition becomes

$$\nu_r(s)/\mu_r(s) = \gamma_r. \qquad (2.18)$$

Formula (2.18) and Equation (2.16) make it clear that γ_r is the risk premium in the sense previously described.

Our final reformulation of the short-rate Primitive Theory is designed to make it look like what we call the "HL/HJM Theory". This theory is presented in the papers [HL] and [HJM], and is discussed in [C 1988]. For this reformulation we put

$$\left.\begin{aligned}\alpha_r(t,q) &= \frac{\partial}{\partial q}[(q-t)\chi_r(q-t)] \equiv -\frac{\partial}{\partial q}[\nu_r(q-t) - \tfrac{1}{2}\mu_r(q-t)^2],\\ \sigma_r(t,q) &= \frac{\partial}{\partial q}[(q-t)\Psi_r(q-t)] \equiv -\frac{\partial}{\partial q}[\mu_r(q-t)].\end{aligned}\right\} \qquad (2.19)$$

(Note that the "≡" in (2.19) follows by (2.17).)

Then we can rewrite Equation (2.60) or (2.14), (2.15), (2.16) as

$$P(s,q) = P(t,q)/_{P(t,s)}\, H(t,s,q) \tag{2.20}$$

where

$$H(t,s,q) = \exp[-\int_{\tau=s}^{q}\int_{\rho=t}^{s}[\alpha_{r_\rho}(\rho,\tau)d\rho + \sigma_{r_\rho}(\rho,\tau)dB_\rho]d\tau \tag{2.21}$$

or equivalently as

$$P(t,q) = \exp[-\int_{\tau=t}^{q} f(t,\tau)d\tau] \tag{2.22}$$

where

$$df(\rho,\tau) = \alpha_{r_\rho}(\rho,\tau)d\rho + \sigma_{r_\rho}(\rho,\tau)dB_\rho. \tag{2.23}$$

Now, (2.22), (2.23) is the starting point of the paper [HJM]. It is natural because f is the "instantaneous forward rate", defined by the following equation, which is equivalent to (2.22):

$$f(t,q) = -\frac{\partial P(t,q)}{\partial q}/P(t,q). \tag{2.24}$$

The perspective of [HJM] is thus to regard the term structure as being driven by the instantaneous forward rates, which obey Equation (2.23). Note that in [HJM] the coefficients α and σ can be more general than those given by Equation (2.19). Therefore the theory in [HJM] is more general than the short-rate Primitive Theory. However the coefficient set χ, Ψ or equivalently μ, ν or $\tilde{\chi}$, $\tilde{\Psi}$ of the Primitive Theory is easier to characterise and to find empirically than the coefficient set α, σ of the general [HJM] Theory.

The Equations (2.20), (2.21) are the continuous-time version of the theory of [HL], as explained in [C 1988].

The advantage of the HL/HJM reformulation of the Primitive Theory is that it is integrated over time. By this we mean that starting from a current time t, it gives the behaviour conditioned on the current term structure, not only at a short time t+ε into the future, but at any time s beyond t. This will be important when we use the Primitive Theory to price at time t, contingent claims on the term structure which mature at time s. In particular we see from (2.23) that the future short rate, conditioned on the term structure at time t, obeys the equation

$$r_q \equiv f(q,q) = f(t,q) + \int_{\rho=t}^{q} \alpha_{r_\rho}(\rho,q)d\rho + \int_{\rho=t}^{q} \sigma_{r_\rho}(\rho,q)dB_\rho. \qquad (2.25)$$

3. HIGHER FACTOR VERSIONS OF THE PRIMITIVE THEORY

The 1-factor versions of the Primitive Theory which is presented in Section 2, cannot be adequate to model the evolution of the term structure, because it is empirically clear that there is more than one dimension of random input in this evolution. This is clear from the empirical fact that the short rate r_t and long rate ℓ_t are not perfectly correlated in their evolution. (The long rate ℓ_t is just the limit for large q of the rate p(t,q).) In fact the papers [SS], [S 1989c] conclude that the long rate ℓ_t and the spread $\ell_t - r_t$ follow orthogonal mean-reverting Brownian motions.

We can modify the Primitive Theory of Section 2 so that it is driven by the rates r_t and ℓ_t together: we refer to this as the 2-factor long-short rate version of the theory. In general we can write the equation for the combined long-short rate as the following analogue of Equaiton (2.4):

$$\left.\begin{aligned} d\ell_t &= \xi_1(\ell_t,r_t)dt+\eta_{11}(\ell_t,r_t)dB_t^1+\eta_{12}(\ell_t,r_t)dB_t^2 \\ dr_t &= \xi_2(\ell_t,r_t)dt+\eta_{21}(\ell_t,r_t)dB_t^1+\eta_{22}(\ell_t,r_t)dB_t^2 \end{aligned}\right\} \tag{3.1}$$

where B_t^1 and B_t^2 are orthogonal (statistically indepenent) Brownian motions. Then we can write as an analogue of (2.6)

$$\begin{aligned} p(t+\varepsilon,q) &= g(t,t+\varepsilon,q) + \chi_{r_t,\ell_t}(q-t)\Delta_t^{t+\varepsilon}t+ \\ &\Psi^1_{r_t,\ell_t}(q-t)\Delta_t^{t+\varepsilon}B^1 + \Psi^2_{r_t,\ell_t}(q-t)\Delta_t^{t+\varepsilon}B^2. \end{aligned} \tag{3.2}$$

Also as an analogue of (2.9) we have

$$\begin{aligned} p(t+\varepsilon,q) &= g(t,t+\varepsilon,q) + \tilde{\chi}_{r_t,\ell_t}(q-t)\Delta_t^{t+\varepsilon}t \\ &+ \tilde{\Psi}^r_{r_t,\ell_t}(q-t)\Delta_t^{t+\varepsilon}r + \tilde{\Psi}^\ell_{r_t,\ell_t}(q-t)\Delta_t^{t+\varepsilon}\ell, \end{aligned} \tag{3.3}$$

and the coefficients $\tilde{\chi}$, $\tilde{\Psi}^r$, $\tilde{\Psi}^\ell$ satisfy the conditions

$$\left.\begin{aligned} &\tilde{\chi}(0) = \lim_{s\to\infty} \tilde{\chi}(s) = 0, \\ &\tilde{\Psi}^r(0) = 1, \lim_{s\to\infty} \tilde{\Psi}^r(s) = 0, \\ &\tilde{\Psi}^\ell(0) = 0, \lim_{s\to\infty} \tilde{\Psi}^\ell(s) = 1. \end{aligned}\right\} \tag{3.4}$$

(One can see condition (3.4) directly by considering Equation (2.13), and the fact that

$\lim_{q\to\infty} p(t+\varepsilon,q) - g(t,t+\varepsilon,q) = \Delta_t^{t+\varepsilon}\ell$.)

The interplay between the original formulation and the "tilda" formulation of the long-short 2-factor theory represented by Equations (3.2) and (3.3) respectively, is basically the same as the interplay for the theory of Section 2, between (2.6) and (2.9). To summarise this:

The original formulation is appropriate for applications of the theory, but the tilda formulation is appropriate for estimation of the theory. In particular, it is reasonable to assume that the coefficients $\tilde{\chi}$, $\tilde{\Psi}^\ell$, $\tilde{\Psi}^r$ are independent of the short rate, and to estimate them by linear regression, using Equation (3.3). Having done this, the coefficients χ, Ψ^1, Ψ^2 can be calculated from $\tilde{\chi}$, $\tilde{\Psi}^\ell$, $\tilde{\Psi}^r$ using a formula analogous to (2.11), which is obtained by substituting for $(\Delta_t^{t+\varepsilon}\ell, \Delta_t^{t+\varepsilon}r)$ in (3.3), using (3.1) in fattened-up form. This formula will also involve the coefficients ξ and η of Equation (3.1), and these must be estimated separately. Note that this estimation is made easier if we assume that the long rate and the spread between long and short rates are orthogonal mean reverting Brownian Motions.

Going beyond Equations (3.1), (3.2), it is clear that a version of the Primitive Theory can be constructed based on any collection of parameters which are taken as driving the term structure. If one writes these parameters as $(x_1,\dots,x_m)$, then one can write an equation giving their evolution as a joint stochastic process, and then write

$$p(t+\varepsilon,q) = g(t,t+\varepsilon,q) + \chi(q-t)\Delta_t^{t+\varepsilon}t + \Psi^1(q-t)\Delta_t^{t+\varepsilon}B^1 + \dots + \Psi^n(q-t)\Delta_t^{t+\varepsilon}B^n \tag{3.5}$$

Also analogous to Equation (2.8) or equivalently (2.18), there is a risk premium for each dimension of randomness in the evolution Equation (3.5), and this gives an equation which can express the drift coefficient χ in terms of the noise coefficients $\Psi^1,\dots,\Psi^n$ in (3.5).

The technique of Factor Analysis is decribed in [S 1989b], and it yields a set of parameters which drive the term structure in an optimal way. These parameters are not necessarily identifiable in economic terms, but they are optimal in the sense that they provide the most efficient way to account for the evaluation of the term structure by a given number of parameters, if errors are measured in a least-squares sense. We will describe the technique of Factor Analysis, tailored to suit the Primitive Theory:

First, restrict attention to the discrete time grid $\{t_0 \le t_1 \le t_2 \le \dots\}$, in which $t_{i+1} - t_i = \varepsilon$ for all i and ε is small, say ε is one month. Next, consider the following collection of time series, each corresponding to a given time to maturity, or "term":

$$\Delta_{t_0}^{t_1} p(t_1), \quad \Delta_{t_1}^{t_2} p(t_2), \quad \Delta_{t_2}^{t_3} p(t_3), \dots \qquad \text{(term 0)}$$

$$\Delta_{t_0}^{t_1} p(t_2), \quad \Delta_{t_1}^{t_2} p(t_3), \quad \Delta_{t_2}^{t_3} p(t_4), \dots \qquad \text{(term } \varepsilon\text{)}$$

$$\vdots \qquad \vdots \qquad \vdots$$

$$\Delta_{t_0}^{t_1} p(t_n), \quad \Delta_{t_1}^{t_2} p(t_{n+1}), \Delta_{t_2}^{t_3} p(t_{n+2}), \dots \qquad \text{(term } n\varepsilon\text{)}$$

(Recall that $\Delta_{t_i}^{t_{i+1}} p(t_j)) \equiv p(t_{i+1},t_j) - g(t_i, t_{i+1},t_j)$, and that the first of these series is just

$\Delta_{t_0}^{t_1}r, \Delta_{t_1}^{t_2}r, \ldots$, by Equation (2.13).) Now form the matrix $C \equiv \{c_{ij}\}$ $i,j = 1,\ldots,n$, in which the entry c_{ij} is the covariance between the series corresponding to terms i and j. There, c_{ij} tells us to what extent the price changes of the time to maturity t_i bond and the time to maturity t_j bond are tied together. Having done this, calculate the eigenvalues $\lambda_1, \lambda_2, \ldots$ of the matrix C and the corresponding eigenvectors, which we will suggestively denote by $(\Psi^1)^2, (\Psi^2)^2, \ldots$. This is not a difficult calculation because the matrix is symmetric; and iterated application of the Rayleigh-Ritz procedure will calculate them in turn, with decreasing eigenvalue, and the iteration can be stopped when the eigenvalue becomes too small to be interesting. This procedure can also give an orthogonal set of eigenvectors. The principal factors yielded by this Factor Analysis are those eigenvectors corresponding to a given number (say m) of the largest eigenvalues, and one can say that they account for a fraction $(\lambda_1 + \ldots + \lambda_m)/(\lambda_1 + \ldots + \lambda_n)$ of the term structure evolution. (n.b. To calculate the sum of all the eigenvectors, i.e. $\lambda_1 + \ldots \lambda_n$, it is not necessary to calculate them individually; this sum is jsut the trace of the matrix C, i.e. $c_{11} + \ldots + c_{nn}$.)

To represent the term structure evolution in terms of the principal factors, just take the noise coefficients in Equation (3.5) to come from these eigenvectors. Note that the drift coefficient χ in Equation (3.5) must be estimated separately. Also, the drift coefficient must depend on the value of the short rate and other factors, in order to be consistent with these factors being mean reverting (recall consistency condition (2.7)), but the noise coefficients as given by factor analysis cannot depend on the values of these factors. Actually, when using the Primitive Theory to value options as in Section 4 below, it is not necessary to know the drift coefficient.

4. VALUING INTEREST RATE OPTIONS VIA THE PRIMITIVE THEORY

In this section we discuss the valuation of contingent claims on the term structure, using the Primitive Theory. As we have already said, for simplicity we will restrict attention to the 1-factor short-rate version of the theory, and we will concentrate on interest rate caps and floors.

We base our valuation on the Martingale Formualtion of [HP], [HK], which deals comprehensively with contingent claims. Thus we take as our basic assets the discount bonds which mature at the various times q and which have value P(t,q) at time t, and the "accumulator", which yields the short rate of interest, and whose value at time t is say $P_0(t)$. The Martingale Formulation provides us with the following two results, which we first state abstractly, and then explain:

PROPOSITION 1: If there are no arbitrage opportunities among the basic assets, then there is a measure $\tilde{P}$ in the space of the Brownian Motion B_t which drives the evaluation of the term structure, such that if this Brownian Motion is governed by $\tilde{P}$, then the basic asset prices are martingales when discounted at the short rate. □

PROPOSITION 2: Consider a claim which is contingent on the short rate and which matures at time q, giving a payoff of φ(r) if the short rate at that time is r. Assume that the claim is attainable, i.e. its payoff can be replicated by a dynamic portfolio comprised of the basic assets. Then the value φ_t of the contingent claim at time t is given by

$$\varphi_t = \tilde{\mathbb{E}}_t [\exp\{-\int_{\tau=t}^{q} r_\tau \, d\tau\} \, \varphi(r_q)], \tag{4.1}$$

where $\tilde{\mathbb{E}}_t$ is the expectation relating to the measure $\tilde{P}$ of Proposition 1. □

The difficulty in Propositions 1 and 2 is to understand the "martingale measure" $\tilde{P}$. This can be characterised in terms of the Girsanov transformation

$$d\tilde{B}_t = dB_t - \gamma_{r_t} dt, \tag{4.2}$$

where $\gamma_r \equiv \nu_r/\mu_r$ is the risk premium of Equation (2.18). To obtain the probability distribution of any asset price process having switched the governing measure to $\tilde{P}$, we can take the stochastic equaiton for the process, and substitute $d\tilde{B}_t$ for dB_t, as given by Equaiton (4.2). Thus, the $\tilde{P}$–distribution of the price P(t,q) of the time q maturity bond is like the distribution of the process $\tilde{P}$(t,q) given by

$$d\tilde{P}(t,q)/\tilde{P}(t,q) = [\nu_{r_t}(q-t) + r_{tp}]\,dt + \mu_{r_t}(q-t)d\tilde{B}_t \tag{4.3}$$

(Cf. Equation (2.16)), and this yields using (4.2),

$$d\tilde{P}(t,q)/\tilde{P}(t,q) = r_t\,dt + \mu_{r_t}(q-t)dB_t. \tag{4.4}$$

From (4.4) it is clear that $\tilde{P}$(t,q) discounted at the short rate is a martingale, as Proposition 1 says it should be, and in fact this can serve as the justification for Equation (4.2).

In order to use Equation (4.1) for valuing contingent claims, we must know the distribution of the short rate governed by the measure $\tilde{P}$. This is given by the following result:

<u>PROPOSITION 3</u>: The distribution of the short rate r_q governed by the measure $\tilde{P}$, and given a knowledge of the term structure at time t, is the same as the distribution of $\tilde{r}_q$, where

$$\tilde{r}_q = f(t,q) + \tfrac{1}{2}\frac{\partial}{\partial q}\int_{\rho=t}^{q} \mu_{\tilde{r}_\rho}(q-\rho)^2 d\rho - \frac{\partial}{\partial q}\int_{\rho=t}^{q} \mu_{\tilde{r}_\rho}(q-\rho)dB_\rho \tag{4.5}$$

$$\equiv f(t,q) + \tfrac{1}{2}\int_{\rho=t}^{q} \frac{\partial}{\partial q}[\mu_{\tilde{r}_\rho}(q-\rho)^2]d\rho - \int_{\rho=t}^{q} \frac{\partial}{\partial q}[\mu_{\tilde{r}_\rho}(q-\rho)]dB_\rho \tag{4.6}$$

PROOF: By Proposition 1, we can obtain this distribution by substituting $d\tilde{B}_\rho$ for dB_ρ in Equation (2.25). Thus

$$\tilde{r}_q = f(t,q) + \int_{\rho=t}^{q} \alpha_{\tilde{r}_\rho}(\rho,q)d\rho + \int_{\rho=t}^{q} \sigma_{\tilde{r}_\rho}(\rho,q)d\tilde{B}_\rho$$

$$= f(t,q) + \int_{\rho=t}^{q} \frac{\partial}{\partial q}[\nu_{\tilde{r}_\rho}(q-\rho) - \tfrac{1}{2}\mu_{\tilde{r}_\rho}(q-\rho)^2]d\rho$$

$$- \int_{\rho=t}^{q} \frac{\partial}{\partial q}[\mu_{\tilde{r}_\rho}(q-\rho)]d\tilde{B}_\rho$$

(using (2.19)), and (4.5) follows by taking the differentials from under the integrals, and using (4.2). The justification for swapping the differential and integral is the fact that $\nu(0) = \mu(0) = 0$, which follows form (2.17), and the general formulae

$$\frac{\partial}{\partial q}\int_{\rho=t}^{q} \Theta(\rho,q)d\rho = \Theta(q,q) + \int_{\rho=t}^{q} \frac{\partial}{\partial q}\Theta(\rho,q)d\rho,$$

$$\frac{\partial}{\partial q}\int_{\rho=t}^{q} \Theta(\rho,q)dB_\rho = \Theta(q,q)\frac{dB_q}{dq} + \int_{\rho=t}^{q} \Theta(\rho,q)dB_\rho.$$

Equation (4.6) follows from (4.5) by putting the differentials back inside the integrals. □

PROPOSITION 4: The contingent claim value in Proposition 2 is given by

$$\varphi_t = \mathbb{E}[\exp\{-\int_{\rho=t}^{q} \tilde{r}_\rho\, d\rho\}\, \varphi(\tilde{r}_q)] \qquad (4.7)$$

where $\tilde{r}_\rho$ is given by Equation (4.5) or (4.6), and $\mathbb{E}$ is the usual expectation.

PROOF: Apply the Girsanov transformation to (4.1). □

Our procedure for valuing the contingent claim of Propositions 2 and 4 is to build up the joint probability density function of the short rate $\tilde{r}_q$ and the discount factor $\exp\{-\int_{\rho=t}^{q} \tilde{r}_\rho \, d\rho\}$, and then to integrate the discounted payoff function over this density function, as in formula (4.7). This density function must be built up as time evolves from t to q, and using the Equation (4.6) for $\tilde{r}_q$, and the following, which is obtained by substituting for $\tilde{r}_\rho$ using (4.5):

$$\exp\{-\int_{\rho=t}^{q} \tilde{r}_\rho \, d\rho\} = P(t,q) \exp\{\int_{\rho=t}^{q} \mu_{\tilde{r}_\rho}(q-\rho) dB_\rho - \tfrac{1}{2} \int_{\rho=t}^{q} \mu_{\tilde{r}_\rho}(q-\rho) d\rho\} \qquad (4.8)$$

Note that the term "exp{ ···}" on the RHS of (4.8) is an exponential martingale, therefore if the contingent claim is just the time q maturity bond, so that $\varphi(\tilde{r}_q) \equiv 1$, then the formula (4.7) gives the correct value for this bond at time t, namely P(t,q).

Proposition 4 tells us how to value an option on the short rate, which pays off only at time q. But a cap or a floor is an option on the 3 month rate, which pays off after every 3 month period up to its maturity time, if the 3 month rate is higher or lower respectively, than the strike rate. Therefore in our evaluation of the cap or floor we should add together the discounted payoffs at 3 monthly intervals, as we build up the density funciton for $\tilde{r}_\rho$ and the discount factor. Also, we should use the formula of the following proposition for the 3 month rate process. This formula is less temperamental than (4.6), because, like (4.8), it does not require derivatives of μ.

Notice that Proposition 5 refers to the continuously compounded rate, whereas a cap or floor is concerned with the simple 3 month rate. However, over 3 months this discrepancy is negligably small.

PROPOSITION 5: Denote by r_q^δ the interest rate at time q, to run over the period [q,q+δ]. Thus, $r_q^\delta = -1/\delta \log P(q,q+\delta)$. Then given the term structure at time t, this process has the same distribution as $\tilde{r}_q^\delta$, which obeys the equation

$$\tilde{r}_q^\delta = g(t,q,q+\delta) + \int_{\rho=t}^{q} \left\{ \frac{\mu_{\tilde{r}_\rho}(q+\delta-\rho)^2 - \mu_{\tilde{r}_\rho}(q-\rho)^2}{\delta} \right\} d\rho$$

$$- \int_{\rho=t}^{q} \left\{ \frac{\mu_{\tilde{r}_\rho}(q+\delta-\rho) - \mu_{\tilde{r}_\rho}(q-\rho)}{\delta} \right\} dB_\rho \qquad (4.9)$$

(n.b. Formula (4.9) is clearly true in the limit for small δ, by Formula (4.6). However we need Formula (4.9) for δ = 3 months, which cannot be considered small in this sense.)

PROOF: Start from the formula

$$f(q,\tau) = f(t,\tau) + \int_{\rho=t}^{q} [\alpha_{r_\rho}(\rho,\tau)d\rho + \sigma_{r_\rho}(\rho,\tau)dB_\rho].$$

Substitute for α, σ in terms of μ, ν using (2.19) as in the proof of Proposition 3; also substitute $d\tilde{B}_\rho$ for dB_ρ using (4.2); also pull out the differentiations, to obtain

$$\tilde{f}(q,\tau) = \tilde{f}(t,\tau) + \frac{\partial}{\partial\tau} \int_{\rho=t}^{q} [\mu_{r_\rho}(\tau-\rho)^2\, d\rho - \mu_{r_\rho}(\tau-\rho)dB_\rho].$$

(Tilda on f indicates that it is governed by the probability $\tilde{P}$.) Now operate on this last equation with $\int_{\tau=q}^{q+\delta} \cdots d\tau$, and the result follows by (2.22). □

REFERENCES

[C 1988] Carverhill, A.P., The Ho-Lee Term Structure Theory: a Continuous Time Version, FORC Preprint 88/5, University of Warwick 1988.

[C 1989] Carverhill, A.P., A Primitive Theory of the Term Structure of Interest Rates, FORC Working Paper, Univeristy of Warwick 1989.

[D] Dybvig, P.H., Bond and Bond Option Pricing Based on the Current Term Structure, Working Paper, Washington Univeristy, St. Louis, Missouri, 1989.

[HK] Harrison, J.M., and Kreps, D.M., Martingales and Arbitrage in Multiperiod Securities Markets, Journal of Economic Theory 20, pp. 381-408, 1979.

[HP] Harrison, J.M. and Pliska S.R., Martingales and Stochastic Integrals in the Theory of Continuous Trading, Stoch. Proc. and Apps. 11, pp. 215-260, 1981.

[HJM] Heath, D., Jarrow, R. and Morton, A., Bond Pricing and the Term Structure of Interest Rates, Preprint, Cornell Univeristy, Ithaca, New York 14853, 1987.

[HL] Ho, T.S.Y., and Lee, S., Term Structure Movements and Pricing Interest Rate Contingent Claims, Journal of Finance 41, pp. 1011-1029, 1986.

[SS] Schaefer, S.M. and Schwartz, E.S., A Two Factor Model of the Term Structure: an Approximate Analytical Solution. JFQA, Vol. 19, pp. 413-424, 1984.

[S 1989a] Steeley, J.M., Estimating the Gilt-Edged Term Structure: Basis Splines and Confidence Intervals. Journal of Business Finance and Accounting 1989 (forthcoming).

[S 1989b] Steeley, J.M., Common Factors Affecting Gilt-Edged Securities, Working Paper, University of Warwick, 1989.

[S 1989c] Steeley J.M., Modelling the Gilt-Edged Term Structure, FORC Preprint 89/2, University of Warwick, 1989.