Financial Options Research Centre

University of Warwick

A Primitive Theory of the Term Structure of Interest Rates

Andrew Carverhill

August 1989

FORC Preprint: 89/4

1. Introduction

Among theories of the term structure of interest rates we distinguish two approaches: that of Cox, Ingersoll and Ross ([CIR]) and Vasicek ([V]), which we identify together and refer to as the "CIR/V theory", and that of Ho and Lee ([HL]) and Heath, Jarrow and Morton ([HJM]), which we refer to as the "HL/HJM theory". See also [C]. Both theories assume that there is just one random factor which drives the term structure, but they differ fundamentally in that the CIR/V theory assumes that the whole of the term structure at a given time is determined by the short rate at that time, and the HL/HJM theory studies the evolution of the term structure from an arbitrary initial structure. These fundamental differences give rise to the contrasting strengths and weaknesses of the two theories; HL/HJM seems preferable in that it is empirically clear that the term structure evolves smoothly but that it does not depend only on the short rate, also CIR/V seems preferable in that it is transparent, and gives a sensible prediction for the long-run behaviour of the term structure, whereas the HL/HJM can easily lead to absurd results in the long-run, such as predicting negative or unbounded interest rates (see [C], [HJM].)

Our aim in this paper is to present a "primitive" term structure theory, which captures the attractive features of both the CIR/V and HL/HJM theories, and yet is not too primitive to be empirically meaningful. The basic assumption of the primitive theory is simply that as time evolves from t to $t+\varepsilon$ (with ε small), then the difference between the forward price at time t, to run from $t+\varepsilon$ and mature at time q, and the spot price at time $t+\varepsilon$ to mature at time q, is perfectly correlated with the evolution of the short rate between times t and $t+\varepsilon$. This assumption actually subsumes that of the CIR/V theory, and is very much like that of HL/HJM.

We present our primitive theory in Section 2 below, and we show how it relates to the CIR/V theory in Section 3, and to the HL/HJM theory in Section 4. In Section 5 we develop the Primitive Theory in the light of empirical estimates of the term structure, and in Section 6 we give a summary and conclusions.

I would like to take this opportunity to thank my colleagues Stewart Hodges, Nick Webber and Jim Steeley for their help in clarifying the ideas presented here; thanks also to Les Clewlow, who did the computer work, and Philip Dybvig, who gave us a preview of his paper [D].

2. The Primitive Theory

Assumptions of the Primitive Theory

In this theory we take the short rate r_t to be an autonomous stochastic process, drive according to the equation

$$dr_t = \zeta(r_t)\,dt + \eta(r_t)dB_t, \tag{1}$$

where the coefficients ζ and η can depend on the short rate, and must be chosen using empirical considerations, and B_t is a standard Brownian Motion. Our basic assumption is formulated as

$$p(t+\varepsilon, q) = g(t, t+\varepsilon, q) + \chi_{r_t}(q-t)\,\Delta_t^{t+\varepsilon} t + \psi_{r_t}(q-t)\,\Delta_t^{t+\varepsilon} B + o_E(\varepsilon) \tag{2}$$

where our notation is as follows:

By $p(s, q)$ we mean the spot rate at time s to mature at time q, i.e $p(s, q) = \frac{-1}{q-s} \log P(s, q)$, is the price at time s of a pure discount bond which yields 1 at time q.

By $g(t, s, q)$ we mean $\frac{-1}{q-s}$ log $G(t, s, q)$, where $G(t, s, q)$ is the forward price given by $P(s, q)/P(t, s)$.

By $\Delta_t^{t+\varepsilon}B$ we mean the increment $B_{t+\varepsilon} - B_t$, so that it is $N(O, t)$, i.e. normally distributed with 0 and variance t.

By $\Delta_t^{t+\varepsilon}t$ we simply mean $(t+\varepsilon) - t \equiv \varepsilon$.

The functions $\chi_r(s)$ and $\psi_r(s)$ must be chosen using the empirical and theoretical considerations discussed below; the fact that their arguments in (2) are both $q - t$ reflects the fact that the model is time homogeneous.

The notation $o_E(\varepsilon)$ represents a remainder term which is of small order in ε in an L^2 sense. This means that if the remainder is X_ε, then $E[X_\varepsilon^2]^{\frac{1}{2}} = o(\varepsilon)$ as $\varepsilon \to 0$.

<u>Reformulation of the Basic Assumption in Terms of Prices Rather than Rates</u>

We can reformulate our basic assumption, equation (2) as any of the following three equations:

$$P(t + \varepsilon, q) = G(t, t + \varepsilon, q)\ [1 + \upsilon_{r_t}(q - t)\ \Delta_t^{t+\varepsilon} t + \mu_{r_t}(q - t)\ \Delta_t^{t+\varepsilon} B] + o_E(\varepsilon); \qquad (3)$$

(Recall: $G(t, t + \varepsilon, q) = P(t, q)/P(t, t + \varepsilon)$)

$$P(t + \varepsilon, q)/P(t, q) = 1 + r_t \Delta_t^{t+\varepsilon} t + \upsilon_{r_t}(q - t)\ \Delta_t^{t+\varepsilon} + \mu_{r_t}(q - t)\ \Delta_t^{t+\varepsilon} B + o_E(\varepsilon)\ ; \qquad (4)$$

$$\frac{dP(t, q)}{P(t, q)} = \upsilon_{r_t}(q - t)dt + r_t\, dt + \mu_{r_t}(q - t)\ dB_t\ . \qquad (5)$$

In (3), (4), (5) the functions υ and μ correspond to χ and ψ of (2), and they are related via

$$\upsilon_r(s) = -s\, \chi_r(s) + \frac{1}{2} s^2\ \psi_r(s)^2\ , \qquad (6)$$

$$\mu_r(s) = -s\psi_r(s) .$$

Proofs of the Reformulations

(2) <=> (3)

Rewriting (2) using $P(s, q) = \exp[-(q - s)\, p(s, q)]$, it becomes

$$P(t + \varepsilon, q) = [P(t, q)/P(t, t + \varepsilon)] \exp[-(q - (t + \varepsilon)\ \{\chi_{r_t}(q - t)\, \Delta_t^{t+\varepsilon} t$$
$$+ \psi_{r_t}(q - t)\, \Delta_t^{t+\varepsilon}B + o_E(\varepsilon)\}]$$

$$= [P(t, q)/P(t, t + \varepsilon)]\ [1 - (q - (t + \varepsilon))\ \{\ --\ \} + \frac{1}{2}(q - (t + \varepsilon))^2\ \{\ --\ \}^2 + o(\{\ ---\ \}^2)$$

(using $\exp[-x] = 1 - x + {}^1\!/_2\, x^2 + o(x^2)$)

$$= [P(t, q)/P(t, t + \varepsilon)]\ [1 - (q - (t + \varepsilon))\, \chi_{r_t}(q - t)\, \Delta_t^{t+\varepsilon} - (q - (t + \varepsilon))\ \psi_{r_t}(q - t)\, \Delta_t^{t+\varepsilon}B + \frac{1}{2}(q - (t + \varepsilon))^2\ \psi_{r_t}(q - t)^2\, \Delta_t^{t+\varepsilon} t] + o_E(\varepsilon)$$

(expanding the curly brackets, and using $dB_t.dB_t = dt$, $dB_tdt = 0$, $dt.\ dt = 0$, and $\Delta_t^{t+\varepsilon} B.\ \Delta_t^{t+\varepsilon} B = \Delta_t^{t+\varepsilon}t$. Of course the penultimate term in this last expression is an Ito-type term).

We obtain this last expression if we substitute for υ and μ in (3), except for some terms involving $\varepsilon\Delta t$ and $\varepsilon\Delta B$, which come from the $(q - (t + \varepsilon))$ factor. These terms can be absorbed into the $o_E(\varepsilon)$ term.

(3) <=> (4)

For this, note that

$$r_t = \lim_{\varepsilon \to 0} [-{}^1/_\varepsilon \log P(t, t + \varepsilon)]$$

$$= \left. \frac{-\partial(\log P(t, t + \varepsilon))/\partial\varepsilon}{\partial\varepsilon/\partial\varepsilon} \right|_{\varepsilon=0}$$

(L'Hopital's Rule)

$$= \left. \frac{-\partial P(t, t + \varepsilon)}{\partial\varepsilon} \right|_{\varepsilon=0}$$

and so

$$P(t, t + \varepsilon)^{\pm 1} = 1 \pm \varepsilon r_t + o(\varepsilon) \ .$$

Substituting for $P(t, t + \varepsilon)^{-1}$ in (3) using this, and multiplying out, yields (4).

(4) <=> (5).

Equation (5) is just a reformulation of (4) in terms of the Ito differential, if we rephrase (4) so that its LHS becomes $(P(t + \varepsilon, q) - P(t, q))/P(t, q)$.

Note

Equation (5) is perhaps the most friendly formulation, and as we have just said, equation (4) is equivalent to (5), except that in (4) we have not gone to the differential limit. This explains the presence of the $o_E(\varepsilon)$ term in (4) (and hence in (3) and (2)); an Ito difference $\alpha_t \Delta_t^{t+\varepsilon} B$ tends to the Ito differential $\alpha_t dB_t$ with error $o_E(\varepsilon)$, or to be more technically correct (since the Ito differential is strictly only meaningful when integrated), the approximate integral

$$\sum_{i=o}^{q-1} \alpha_i \ \Delta_{t_i}^{t_{i+1}} B \ \text{for} \ a = t_o < t_1 < \cdots < t_a = b$$

tends to the integral $\int_{t=a}^{b} \alpha_t dB_t$ in the L^2 norm, i.e. with error $0_E(\varepsilon)$, where

$$\varepsilon = \max_{i = o, \ldots, q-1} \{t_{i+1} - t_i\} \quad .$$

Arbitrage Across Maturities: Constancy of the Risk Premium

We will take the constancy of the risk premium as one of the assumptions of the Primitive theory; it is explained in the following proposition. This idea is common to the CIR/V and HL/HJM theories.

Proposition 1

If our Primitive Theory holds, and there are no arbitrage opportunities, then we must have

$$\frac{\upsilon_r(s)}{\mu_r(s)} \ (\equiv \gamma_r \text{ say}) \tag{7}$$

independent of s . This γ_r is the risk premium; it gives the expected return in excess of the riskless return, which is yielded when the bond has "risk" μ, i.e. $\upsilon = \gamma\mu + r$. Note that γ can depend on the current short rate.

Proof

At time t form a portfolio comprising quantities ξ_1 and ξ_2 of discount bonds to mature at times q_1 and q_2 respectively. Then the value of this portfolio at time t is

$$\xi_1 P(t, q_1) + \xi_2 P(t, q_2) , \tag{8}$$

and at time $t + \varepsilon$ the value is

$$\xi_1 P(t + \varepsilon, q_1) + \xi_2 P(t + \varepsilon, q_2) ,$$

which (using (3)) is

$$\xi_1 [P(t , q_1)/P(t, t + \varepsilon)] [1 + \upsilon_{r_t}(q_1 - t) \Delta_\varepsilon^{t+\varepsilon} t + \mu_{r_t}(q_1 - t) \Delta_t^{t+\varepsilon} B]$$

$$+ \xi_2 [P(t , q_2)/P(t, t + \varepsilon)] [1 + \upsilon_{r_t}(q_2 - t) \Delta_\varepsilon^{t+\varepsilon} t + \mu_{r_t}(q_2 - t) \Delta_t^{t+\varepsilon} B]$$

$$+ o_E(\varepsilon) . \tag{9}$$

Now, suppose we have chosen ξ_1 , ξ_2 so that the portfolio is hedged to order $0_E(\varepsilon)$, i.e. so that the ΔB terms in (9) cancel and we have

$$\xi_1 P(t , q_1)\mu_{r_t}(q_1 - t) + \xi_2 P(t, q_2) \mu_{r_t}(q_2 - t) = 0 . \tag{10}$$

Then to prevent arbitrage possibilities, we must have

$$\frac{\text{value at time } t+\varepsilon, \text{ i.e } (9)}{\text{Value at time } t, \text{ i.e } (8)} = 1/P(t, t+\varepsilon) + o_E(\varepsilon) \ . \tag{11}$$

Substituting into (11) and taking ξ_1 and ξ_2 to satisfy (10), we conclude that

$$\frac{\upsilon_{r_t}(q_1 - t)}{\mu_{r_t}(q_1 - t)} = \frac{\upsilon_{r_t}(q_2 - t)}{\mu_{r_t}(q_2 - t)} \ ,$$

and hence our result follows, since q_1 and q_2 were chosen arbitrarily.

Recovering the Short Rate: A Consistency Condition

From our basic evolution equation (2), we can recover an equation for the short rate, which must agree with equation (1), and this leads to the following result.

Proposition 2

The assumptions of the Primitive Theory imply that

$$\zeta(r) = \chi_r(0) \ , \ \eta(r) = \psi_r(0) \tag{12}$$

Proof

Choose $\varepsilon > 0$. Putting $q = t + 2\varepsilon$ in equation (2), we have

$$P(t+\varepsilon, t+2\,\varepsilon) = \{2\,P(t, t+2\,\varepsilon) - P\,(t, t+\varepsilon)\} + \chi_{r_t}(2\varepsilon)\,\Delta_t^{t+\varepsilon} t + \psi_{r_t}(2\,\varepsilon)\,\Delta_t^{t+\varepsilon} B$$
$$+ o_E(\varepsilon)$$

Now, note that

$$p(t+\varepsilon, t+2\varepsilon) = r_{t+\varepsilon} + o(\varepsilon),$$

$$2p(t, t+2\,\varepsilon) - p(t, t+\varepsilon) = r_t + o(\varepsilon),$$

$$\chi_r(2\varepsilon) = \chi_r(0) + o(\varepsilon), \qquad \text{(assuming } \chi_r(s) \text{ and } \psi_r(s) \text{ are smooth in s)}$$

$$\psi_r(2\varepsilon) = \psi_r(0) + o(\varepsilon),$$

to deduce that

$$r_{t+\varepsilon} - r_t = \chi_{r_t}(0)\,\Delta_t^{t+\varepsilon}t + \psi_{r_t}(0)\,\Delta_t^{t+\varepsilon}B + o_E(\varepsilon)$$

and hence that

$$dr_t = \chi_{r_t}(0)dt + \psi_{r_t}(0)dB_t,$$

which must agree with equation (1).

3. The Primitive Theory vs. the CIR/V Theory

The CIR/V theory makes all the assumptions of the Primitive Theory, and the additional assumption that the entire term structure at a given time is determined by the short rate at that time. These assumptions suffice to determine a Black-Scholes type equation for the term structure, as in the following proposition:

Proposition 3

Under the assumption of the CIR/V theory we can write $P_r(t, q)$ for the price at time t of a discount bond which yields 1 at time q, if the current short rate is r. Also we have the backward diffusion equation

$$\frac{\partial P_r}{\partial t}(t, q) = -\frac{1}{2}\sigma_r^2 \frac{\partial^2 P_r}{\partial r^2}(t, q) + (\gamma_r \sigma_r - s_r)\frac{\partial P_r}{\partial r}(t, q) + rP_r(t, q) \tag{13}$$

with final condition

$$P_r(q, q) \equiv 1 \quad \text{for all } r .$$

<u>Proof</u>

The Ito formula yields

$$dP_{r_t}(t, q) = \frac{\partial P_{r_t}}{\partial t}(t, q)\, dt + \frac{\partial P_{r_t}}{\partial t}(t, q)\, dr_t + \frac{1}{2}\frac{\partial^2}{\partial r^2} P_{r_t}(t, q) \; \langle dr_t , dr_t \rangle$$

$$= \left[\frac{\partial P_{r_t}}{\partial t} + \rho_{r_t}\frac{\partial P_{r_t}}{\partial r} + \frac{1}{2}\sigma_{r_t}^2 \frac{\partial^2}{\partial r^2} P_{r_t}\right] dt + \sigma_{r_t}\frac{\partial P_{r_t}}{\partial r}\, dB_t \tag{14}$$

(the applicability of the Ito formula rests simply on the fact that the process r_t enters as the argument of the function in $P_{rt}(t, q)$.)

Comparing (14) to (5) we obtain

$$\upsilon = 1/p\left[\frac{\partial P}{\partial t} + \rho\frac{\partial P}{\partial r} + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial r^2} P\right], \quad \mu = 1/p\left[\frac{\sigma \partial P}{\partial r}\right], \tag{15}$$

and the equation (13) follows by substituting into (7), using (15).

Notes Concerning the Long Rate

(i) The equation (13) can be solved using standard numerical methods, and both the papers [CIR] and [V] give analytic solutions, for their assumptions about the forms of the coefficients ζ, η and γ. Their analytic solutions are completely different from each other, because their assumptions about ζ, η and γ are different. However, they both predict that as q increases from t, the spot rate $p_{r_t}(t, q)$ will move smoothly from r_t to an asymptotic value (the long rate) which is independent of r_t. This prediction of constant long rates is not unreasonable for a one-factor model.

(ii) In detail, [CIR] assumes that

$$dr_t = \tilde{\kappa}\,(\tilde{\theta} - r_t)\, d_t + \tilde{\sigma}\sqrt{r_t}\; dB_t\,,\ \gamma_r = \tilde{\lambda}\sqrt{r}/\tilde{\sigma} \tag{16}$$

and [V] assumes that

$$d_{r_t} = \tilde{\alpha}\,(\tilde{\gamma} - r_t)\; dt + \tilde{\rho}\; dB_t\,,\ \gamma_r = \tilde{q}\,, \tag{17}$$

where in (16) and (17) the symbols with tildas are all positive constants (and the tildas prevent clashes with other notation.). It is difficult to argue on empirical or theoretical grounds in favour of either of (16) or (17) rather than the other, and yet these differences lead to important differences between the conclusions of [CIR] and [V], and this fact must be regarded as a criticism of the CIR/V theory. In particular, in [CIR] the long rate is

$$\alpha\,\tilde{\kappa}\;\tilde{\theta}/\{[(\tilde{\kappa} + \tilde{\lambda})^2 + 2\,\tilde{\sigma}^2]^{1/2} + [\tilde{\kappa} + \tilde{\lambda}]\} \tag{18}$$

(see [CIR] equation (26)), and in [V] the long rate is

$$\tilde{\gamma} + \widetilde{\rho q}/\tilde{\alpha} - \frac{1}{2}(\tilde{\rho}/\tilde{\alpha})^2 \tag{19}$$

(see [V] equation (28)), and it seems that there is no elegant formulation for the long rate that encompasses (18) and (19) together.

(iii) In the generalised formulation of the CIR/V theory of this present paper, it seems to be unclear whether the long rate is indeed constant.

The following proposition explains how the conclusions of the Primitive Theory can be made to encompass those of the CIR/V theory, by making appropriate choices for the functions μ and υ, or equivalently χ and ψ.

Proposition 4

Suppose we take μ and υ in the Primitive Theory as in equation (15), which relates to the CIR/V theory and with P being the solution to equation (13). Also suppose we start the evolution of the Primitive Theory at time t (i.e. the evolution driven by equation (2), (3), (4) or (5)) from a term structure which is admissible for the CIR/V theory (i.e. with prices $\{P_{r_t}(t, q): q > t\}$, where r_t is the current short rate, and $P_r(t, q)$ is the solution for equation (12)). Then the evolution according to the Primitive Theory will conform to that according to the CIR/V theory.

Proof

The result follows because both theories give the same stochastic equation for $P_{r_t}(t, q)$, namely (5) and (14).

One might say that the difference between the Primitive Theory and the CIR/V Theory is that the Primitive Theory concentrates directly on the functions υ and μ of the equation (5), and does not assume that they are as in equation (14).

4. The Primitive Theory vs the HL/HJM Theory

Formulation of the HL/HJM Theory

The paper [HJM] starts from the formulation

$$P(t, q) = \exp [- \int_{t}^{s} f(t, \tau)\, d\tau], \tag{20}$$

where f is the forward rate, and it satisifes the following stochastic equation for fixed τ and variable $\rho < \tau$:

$$df(\rho, \tau) = \alpha(\rho, \tau)\, d\rho + \sigma(\rho, \tau)\, dB_{\rho} , \tag{21}$$

where B_{ρ} is a standard Brownian notion. This formulation is equivalent to the following, which is presented in [C] as the continuous time analogue of the formulation of [HL]:

For $t < s < q$ we have

$$P(s, q) = [P(t, q)/P(t, s)]\, H(t, s, q) , \tag{22}$$

where

$$H(t, s, q) = \exp\ [- \int_{\tau=s}^{q} \int_{\rho=t}^{s} [\alpha(\rho, \tau) d\rho + \sigma(s, \tau)\, dB_{\rho}]\, d\tau\ . \tag{23}$$

In the spirit of [HL], [HJM], equations (22) and (23) tell us how to predict the term structure at time s given that at time t, and in terms of the random input $\{B_{\rho}\}_{\rho \geq t}^{\rho \leq s}$.

From equations (20) - (23) we see that the HL/HJM theory takes the forward rates of equation (21) to be the fundamental driving process of the term structure, rather than the short rate, and the work of [HL], [HJM] and [C] is mostly concerned with arriving at reasonable choices for the coefficient functions α and σ .

Attempts to Solve the HL/HJM Theory

It is reasonable, as a first attempt to solve the model, to assume that α and σ are nonrandom, i.e. that they do not depend on the behaviour of the term structure. This assumption is essentially made in [HL] itself (compare equation (22) with equations (7) and (8) of [HL]) and in [HJM] Section 7, and in [C]. With this assumption, and applying the arguments of [HL], it is possible to solve the model completely (see [C], also [HJM]), to obtain

$$\sigma(s, \gamma) \equiv \sigma(\text{constant}) \ , \quad \sigma(s, \gamma) = - \gamma \sigma - \sigma^2(\tau-\rho), \tag{24}$$

where γ is the risk premium. (We make no assumptions about the form of γ).

From (21) and (22) we obtain

$$f(s, \tau) = f(t, \tau) - \sigma \int_{\rho=t}^{s} \gamma d\rho + \frac{1}{2} [(\gamma - t)^2 - (\gamma - s)^2] + \sigma[B_s - B_t], \tag{25}$$

and hence

$$r_s \equiv f(s, s) = f(t, s) - \sigma \int_{\rho=t}^{s} \gamma d\rho + \sigma^2/2 \ (s - t)^2 + \sigma(B_s - B_t) \ . \tag{26}$$

Also, we can use (25) in (20) - (23) to obtain explicit formulae for the evolution of the term structure. Unfortunately, these solutions are absurd; for instance (26) predicts that the short rate grows unboundedly, and that with high probability it can become negative.

To avoid such absurd conclusions, we must allow the coefficients α and σ to be random. This is done in [HJM] Section 8, where σ is assumed to be small when the forward rate is small (a device to ensure that the short rate remains positive), but then α must be made random to present arbitrage - across - maturities. If α and σ are random, then it seems that they will be very difficult to interpret and to measure empirically.

The Primitive Theory vs the HL/HJM Theory

To compare these theories, it is useful to rephrase the HL/HJM formulation as

$$P(s, q) = g(t, s, q) + \frac{1}{q-s}\int_{\tau=s}^{q}\int_{\rho=t}^{s}[\alpha(\rho, \tau)\, d\rho + \sigma(\rho, \tau)\, dB_\rho]d\tau.$$

From this we can deduce that

$$p(t + \varepsilon,q) = g(t, t + \varepsilon,q) + \frac{1}{q - t}\left\{\int_{\tau=t}^{q} \alpha(t, \tau)\, d\tau\right\} \Delta_t^{t+\varepsilon} t + \frac{1}{q - t}\left\{\int_{\tau=t}^{q} \sigma(t, \tau)\, d\tau\right\} \Delta_t^{t+\varepsilon} B + o_E(\varepsilon), \tag{27}$$

and comparing (27) with (2) yields

$$\chi_{r_t}(q - t) = \frac{1}{q - t}\int_{\tau=t}^{a} \alpha(t, \tau)\, d\tau \quad , \quad \psi_{r_t}(q - t) = \frac{1}{q - t}\int_{\tau=t}^{q} \sigma(t, \tau) d\tau \tag{28}$$

From (28) we see that we can recast the Primitive formulation into the HL/HJM formulation (i.e. we can go from χ , ψ (or υ, μ) to α , σ), if we allow α and σ to depend on the short rate; infact

$$\alpha_r(t, q) = \frac{d}{dq}[(q - t)\,\chi_r(q - t)]\,, \qquad \sigma_r(t, q) = \frac{d}{dq}[(q - t)\;\psi_r(q - t)]\;. \tag{29}$$

Also if χ and ψ are such that they prevent arbitrage - across - maturities, then so do α and σ of the HL/HJM theory, if they are as in (29). Also, we can go from α , σ to χ , ψ (or v, μ) if $\alpha(\rho, \tau)$ and $\alpha(\rho, \tau)$ depend on r_ρ , but we cannot do this if α and σ are random in a more general way than this. Thus we conclude that the Primitive Theory is a well-motivated special case of the HL/HJM Theory.

Some Technical Notes

(i) Since the HL/HJM formulation given here starts from the forward rates, it is instructive for our comparison between HL/HJM and the Primitive Theory, to derive the equation for forward rates in the Primitive Theory. Note first that from (19) we have

$$f(t, q) = -\frac{\partial P}{\partial q}(t, q)/P(t, q)\;. \tag{30}$$

also from (5) we have

$$P(t, q) = P(0, q)\exp\left\{\int_{\rho=0}^{t}[\upsilon\,(s, q)\,d\rho - r_\rho d\rho + \mu(\rho, q)\,dB_\rho + \mu(\rho, q)^2 d\rho]\right\} \tag{31}$$

and hence

$$\frac{\partial P}{\partial q}(t, q) = \frac{\partial P}{\partial q}(0, q)\exp\left\{\int_{\rho=0}^{t}[\;...]\right\}$$

$$+\,P(0, q)\;\exp\left\{\int_{\rho=0}^{t}[\;...]\right\}\cdot\int_{\rho=0}^{t}\frac{\partial}{\partial q}[\;...\;]$$

$$= P(t, q)\frac{\partial P}{\partial q}(0, q)/\,P(0, q) + P(t, q)\;.\;\int_{\rho=0}^{t}\frac{\partial}{\partial q}[\;...\;] \tag{32}$$

(Substituting using (31).)

Using (30) in (32) we have

$$f(t, q) = f(0, q) - \int_{s=0}^{t} \frac{\partial}{\partial \mu} [\ ...\ ] ,$$

which yields the stochastic equation (for fixed q and time variable t with $t < q$)

$$df(t, q) = -\frac{\partial}{\partial q} \upsilon(t, q)\, dt + \mu(t, q) \frac{\partial}{\partial q} \mu(t, q)\, dt - \mu \frac{\partial \mu}{\partial a} (t, q)\, dB_t. \qquad (33)$$

(ii) Comparing (33) with (21) yields the correspondence

$$\alpha(\rho, \tau) = -\frac{\partial \nu}{\partial \psi} (\rho, \tau) + \mu(\rho, \tau) \frac{\partial \mu}{\partial \tau} (\rho, \tau) ,$$

$$\sigma(\rho, \tau) = \frac{-\partial \mu}{\partial \tau} (\rho, \tau) . \qquad (34)$$

between HL/HJM and the Primitive Theory. This agrees with the correspondence (26) in view of the correspondence (6) between μ, υ and ψ, χ .

(iii) Having obtained equation (33) for the forward rate in the Primitive Theory, we can go on to obtain from it an equation for the short rate, which is like equation (24) of the HL/HJM theory. This equation is

$$r_q \equiv f(q, q) = f(t, q) - \int_{s=t}^{q} \left\{ [\frac{\partial \upsilon}{\partial q} (\rho, q) - \mu(\rho, q) \frac{\partial \mu}{\partial q} (\rho, q)]\ d\rho + \frac{\partial \mu}{\partial q} (\rho, q)\, dB_\rho \right\} \qquad (35)$$

Equation (35) gives the stock price process as time q increases from t , but conditioned on a knowledge of the forward rate curve (or equivalently the term structure itself) at time t . Note that this equation cannot easily be reduced to equation (1) for the unconditioned behaviour of the short rate.

5. Empirical Considerations

This section is intended only as a pilot study, to show how to estimate the coefficients of the Primitive Theory. The application of the theory to interest rate option evaluation will be studied in subsequent work. We believe that for a more detailed empirical study of the Primitive Theory, we should develop a more sophisticated version of the theory, because it is empirically clear that there is more than one dimension of random input in the term structure evolution. Infact the paper [S3] identifies three factors of random input. One might say that the Primitive Theory as developed in this paper gives an optimal account of the evolution of the term structure, in as much as it is driven by the short rate. As our empirical data we use the 4-weekly term structure estimates of [S1]. These refer to British gilts, and cover the period 6 March 1986 to 15 October 1987.

For our empirical study, we cast the Primitive model into another formulation. This formulation involves eliminating the noise term in equation (2) using (1), to obtain

$$p(t + \varepsilon, q) = g(t, t + \varepsilon, q) + \tilde{\chi}_{r_t}(q - t)\, \Delta_t^{t+\varepsilon} t + \tilde{\psi}_{r_t}(q - t)\, \Delta_t^{t+\varepsilon} r + 0_E(\varepsilon), \qquad (36)$$

where

$$\tilde{\chi}_r(q - t) = \tilde{\chi}_r(q - t) - (\zeta(r)/\sigma(r))\, \eta_r(q - t),$$
$$\tilde{\psi}_r(q - t) = (1/\zeta(r))\, \psi_r(q - t). \qquad (37)$$

Equation (36) makes clear the fact that in our theory, the term structure is driven by the short rate.

We will actually estimate $\tilde{\chi}$ and $\tilde{\psi}$. Having done this, χ and ψ can be estimated using (33) if we als estimate the parameters ζ and η of equation (1). The parameters ζ and η are estimated in paper [S2]. We will assume that $\tilde{\chi}$ and $\tilde{\psi}$ are independent of the short rate r . This is perhaps not unreasonable, but it is unreasonable to assume that χ and ψ are independent of r, in view of Proposition 2; if they were, then ζ and η would be also and the short rate would not be mean-reverting.

Our procedure for estimating $\tilde{\chi}$ and $\tilde{\psi}$ is as follows: First denote by $\{t_1, t_2, ...\}$ the times at which the term structures have been calculated. Thus $t_i - t_{i-1}$ is one month for all i. Take ε of equation (36) to be one month. Now for each time to maturity of say k months calculate empirically the quantities

$$\Delta_{t_1}^{t_2} p(t_{1+k}) \quad , \quad \Delta_{t_2}^{t_3} p(t_{2+k}), \; ...$$

$$\Delta_{t_1}^{t_2} r \quad , \quad \Delta_{t_2}^{t_3} r \; , \; , ...$$

where

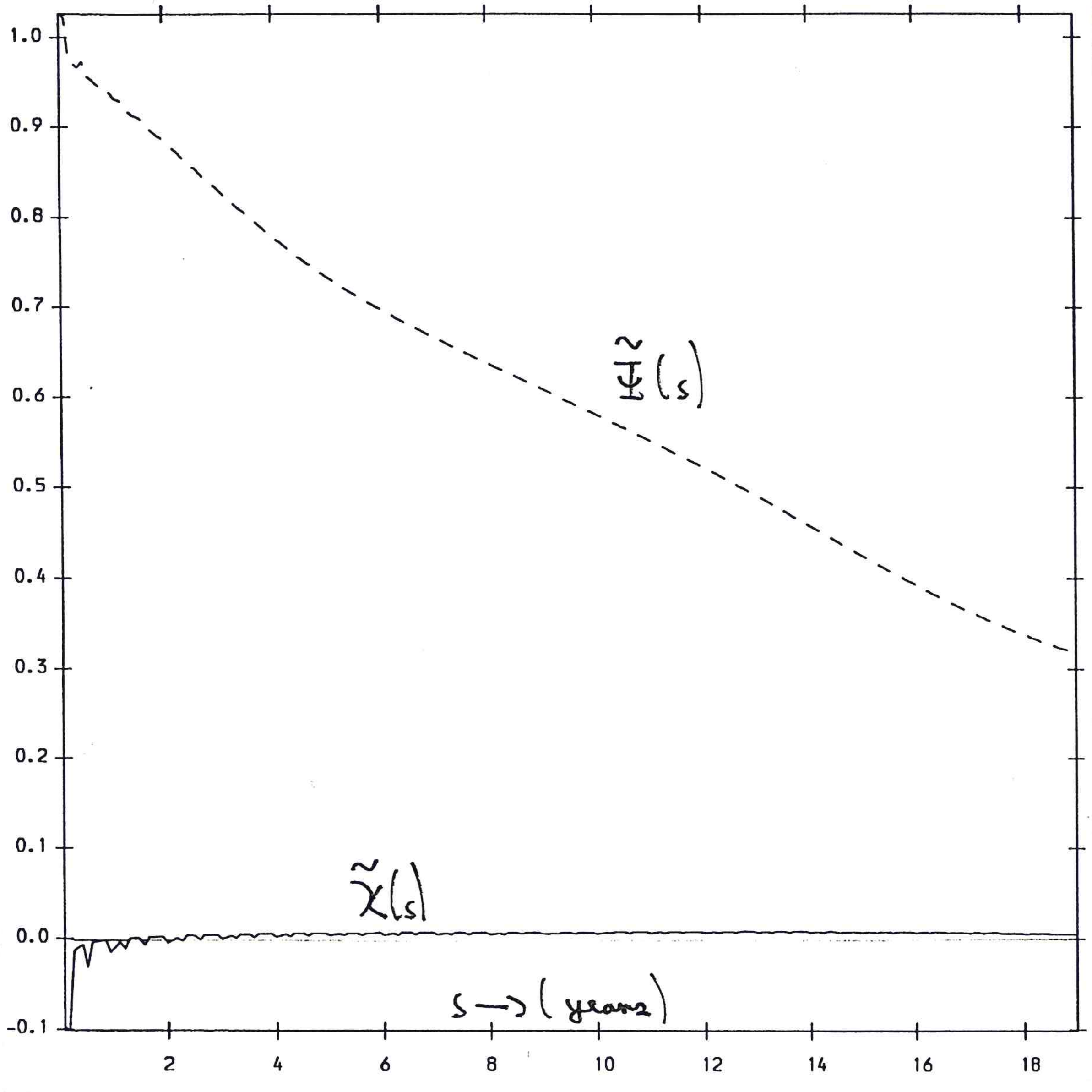
$$\Delta_{t_i}^{t_{i+1}} p(t_{i+k}) = P(t_{i+1}, t_{i+k}) - g(t_i, t_{i+1}, t_{i+k}),$$

and the short rate r_{t_i} is actually calculated as the one month rate $p(t_i, t_{i+1})$.

Finally obtain $\tilde{\chi}$ (k months) and $\tilde{\psi}$ (k months) by a least-square regression on the equation

$$\Delta_{t_i}^{t_{i+1}} p(t_{i+k}) = \tilde{\chi}(\text{k months}). \text{ (1 month)} + \tilde{\psi} \text{ (k months) } \Delta_{t_i}^{t_{i+1}} r \,. \tag{38}$$

The result of this estimation is given in the following graph:



Our estimated $\tilde{\psi}$ in this graph behaves in a reasonable manner; it decreases gradually from 1 and shows that the volability of the short rate is reflected in that of the long rate but is an attenuated manner. Our estimated $\tilde{\chi}$ is more difficult to explain; it seems to be unstable for short term-to-maturity.

We suspect that in a statistical sense $\tilde{\chi}$ is not very robustly determined.

6. Summary and Conclusions

We have presented a Primitive Theory of the evolution of the term structure of interest rates, in which the basic assumption is simply that the evolution of rates for all maturities is perfectly correlated to that of the short rate. We have shown that this theory subsumes the theories of [CIR] and [V]. and is a well-motivated special case of the theory of [HJM]. We have also indicated a procedure for estimating the coefficients of the Primitive Theory.

There is much empirical work to be done to develop and verify the Primitive Theory. Before this is done it will be necessary to develop a higher-factor version of the theory, because it is empirically clear that the short rate alone is not adequate to determine the evolution of the term structure.

REFERENCES

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

[CIR] Cox J. C. Ingersoll J. E. and Ross. S. A. A Theory of the Term Structure of Interest Rates, Econometrica 53, pp. 385-407, 1985.

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

[HJM] Heath, D. Jarrow R. and Morton A., Bond Pricing and the Term Structure of Interest Rates, Preprint, Cornell University, 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.

[S1] Steeley J. M., Estimating the Gilt-Edged Term Structure: Basis Splines and Confidence Intervals. Journal of Business Finance and Accounting 1989 (forthcoming).

[S2] Steeley J. M., Modelling the Gilt-Edged Term Structure, FORC Preprint 89/2, University of Warwick, 1989.

[S3] Steeley J. M., Common Factors Affecting Gilt-Edged Securities, Working Paper, University of Warwick, 1989.

[V] Vasicek O. A. An Equilibrium Characterisation of the Term Structure, Journal of Financial Economics 5, pp. 177-188, 1977.