WHICH MODEL FOR TERM-STRUCTURE OF INTEREST RATES SHOULD ONE USE?[*]

L.C.G. ROGERS[†]

1. Introduction. There are presently many different models of the term structure of interest rates, but little agreement on any one natural one. This is perhaps not surprising in view of the nature and quality of data on term structure; prices of coupon-bearing and zero-coupon bonds, LIBOR rates, index-linked bonds, together with options and futures on such things, all provide information about interest rates, and have to be compatible with any successful model, either by being used to estimate parameters of the model, or by being consistent with the predictions of the (fitted) model. In addition, interest rate swaps provide information, but this is a different market, and should not be expected necessarily to fit the same model.

Much of the work done in this area describes models which one *could* use, rather than claiming strongly that one *should* use them; I have very little to say about what model(s) one *should* use, but will say a few things about what one should *not* use! To decide this, it helps to be clear about what the goal is; we want to build a model that practitioners may rely on. Now a practitioner wants a model which is:

(a) *flexible* enough to cover most situations arising in practice;
(b) *simple* enough that one can compute answers in reasonable time;
(c) *well-specified*, in that required inputs can be observed or estimated;
(d) *realistic*, in that the model will not do silly things.

Additionally, the practitioner shares the view of an econometrician who wants

(e) *a good fit* of the model to data;

and a theoretical economist would also require

(f) *an equilibrium derivation* of the model.

For the practitioner, (a)-(e) already constitute Nirvana!

The fundamental object of study for term-structure is the spot-rate process $(r_t)_{t \geq 0}$. This is a continuous-time process, often assumed to have continuous paths, though sometimes also modelled (more realistically) with jumps. The spot rate represents the instanteneous rate of riskless return

at any time, so $1 invested at time t will have grown by later time T to

$$\$\exp\left(\int_t^T r_u\,du\right).$$

Now if one assumes that r is a stochastic process, the problem of pricing a bond is non-trivial: how much should one pay at time t to receive $1 at later time T? Arbitrage pricing theory says the price must be

$$P(t,T) \equiv E\left[\exp\left(-\int_t^T r_u\,du\right)|\mathcal{F}_t\right]$$

where the expectation is with respect to the "risk-neutral" measure. So the central question is how to model the law of r under the risk-neutral measure so as to achieve (a)-(e).

The plan of the rest of the paper is as follows. Section 2 discusses some very simple Gaussian models for r, and provides a bound for the error arising from the possibility that r can be negative. Section 3 presents some simple one-factor squared-Gaussian models, with an appendix summarizing the analysis of this class of models. Section 4 discusses some two-factor models which have been proposed. In particular, we take a very general class of multi-factor squared Gaussian models and analyse the term-structures which can arise, and in Section 5 we discuss some "whole-yield" models currently in vogue. Section 6 is a discussion of the relationship between equilibrium and no-arbitrage pricing; essentially, the two are the same. Finally, in Section 7, we briefly survey the questions of estimating and fitting with observed data. The relevant literature is surveyed in the corresponding Sections.

As a piece of notation, we define the *yield curve* at time t to be the function

$$T \mapsto -\frac{1}{T-t}\log P(t,T) \quad (T > t)$$

and the *forward rate* f_{tu} (for $0 \le t \le u$) to be given by

$$P(t,T) = \exp\left[-\int_t^T f_{tu}\,du\right],$$

provided the yield curve at time t is differentiable.

2. Simple Gaussian models. If $\alpha, \beta, \sigma : \mathbb{R}^+ \to \mathbb{R}^+$ are any locally bounded functions, then the stochastic differential equation

$$(2.1) \qquad dr_t = \sigma_t\,dW_t + (\alpha_t - \beta_t r_t)dt$$

for r has a unique solution, which is a Gaussian process. This is the model considered by Hull & White [27], generalising the models of Vasicek [42], in which the functions α, β, σ are constant, and of Merton [34] where also $\beta = 0$. The analysis of (2.1) is very simple. Taking

$$K_t \equiv \int_0^t \beta_u \, du,$$

and multiplying (2.1) by $\exp(K_t)$ we obtain

$$d(e^{K_t} r_t) = e^{K_t}(\sigma_t dW_t + \alpha_t dt)$$

from which

$$(2.2) \qquad r_t = e^{-K_t} \left\{ r_0 + \int_0^t e^{K_u}(\sigma_u dW_u + \alpha_u du) \right\}.$$

This is a Gaussian process, for which

$$(2.3) \qquad \mu_t \equiv E r_t = e^{-K_t} \left\{ r_0 + \int_0^t e^{-K_u} \alpha_u du \right\}$$

$$(2.4) \qquad \rho(s,t) \equiv \operatorname{cov}(r_s, r_t) = e^{-K_s - K_t} \int_0^{s \wedge t} e^{2K_u} \sigma_u^2 du.$$

Thus

$$Z_t \equiv \int_0^t r_u \, du \sim N(m_t, v_t)$$

where

$$(2.5) \qquad m_t \equiv \int_0^t e^{-K_u} \left(r_0 + \int_0^u e^{K_s} \alpha_s ds \right) du,$$

$$(2.6) \qquad v_t \equiv 2 \int_0^t du \int_0^u ds \int_0^s dy \sigma_y^2 e^{-K_s - K_u + 2K_y},$$

and hence the bond price is

$$(2.7) \qquad E \exp\left(-\int_0^t r_u \, du \right) = \exp\left(-m_t + \frac{1}{2} v_t \right)$$

(2.8) $= \exp\{-r_0 B(0,t) - A(0,t)\},$

where for $0 \le t \le T$

(2.9) $B(t,T) \equiv \int\limits_{t}^{T} \exp(-K_u + K_t) du,$

(2.10) $A(t,T) \equiv \int\limits_{t}^{T} du \int\limits_{t}^{u} ds \left\{ \alpha_s e^{K_s - K_u} - \int\limits_{t}^{s} dy \sigma_y^2 e^{-K_s - K_u + 2K_y} \right\}.$

Thus this model is extremely *simple*, and the log-Gaussian distribution of bond prices makes it very easy to price derivative securities. For example, the price at time t of a European call option to be exercised at time T with strike X on a zero-coupon bond of maturity $T' > T$ is simply

$$E\left[\exp\left(-\int\limits_{t}^{T} r_u du \right) \{\exp(-r_T B(T,T') - A(T,T')) - X\}^+ \,|\mathcal{F}_t \right]$$

which can be evaluated in closed form. Similarly the value at time t of a futures contract, delivery date T, on a zero-coupon bond with maturity $T' > T$ is

$$E(P(T,T')|F_t) = E\left[\exp(-r_T B(T,T') - A(T,T')) | F_t \right]$$

which can again be evaluated in closed form.

Is the model *well-specified*? Hull & White argue that if one knows at some time $(0$, say$)$ the volatility of r, and the volatility of bonds of all maturities, then since

$$dP(t,T) = -P(t,T)d\left(r_t B(t,T) + A(t,T) \right) + \frac{1}{2} P(t,T) B(t,T)^2 d\langle r\rangle_t,$$

the volatility of the maturity-T bond is $\sigma_0 B(0,T) P(0,T)$; and, since this is known, one can deduce $B(0,T)$ for all T. Then, since $B(0,\cdot)$ is known, we can recover $A(0,\cdot)$ from (2.8). But now knowing $B(0,\cdot)$ we can find K and therefore β from (2.9); and then differentiating $A(0,\cdot)$ gives

$$e^{K_T} A'(0,T) = \int\limits_{0}^{T} \left(\alpha_s e^{K_s} - \int\limits_{0}^{s} \sigma_y^2 e^{2K_y - K_s} dy \right) ds,$$

differentiating once more gives

$$e^{K_T} \left(\beta_T A'(0,T) + A''(0,T) \right) = \alpha_T e^{K_T} - \int\limits_{0}^{T} \sigma_y^2 e^{2K_y - K_T} dy,$$

and another derivative gives

$$e^{2K_T} \left(2\beta_T^2 A'(0,T) + 3\beta_T A''(0,T) + \beta_T' A'(0,T) + A'''(0,T)\right)$$
$$= \left(\alpha_T' + 2\beta_T \alpha_T - \sigma_T^2\right) e^{2K_T}.$$

This cannot be uniquely solved for α, σ, but it gives us an equation

$$\alpha_T' + 2\beta_T \alpha_T - \sigma_T^2 = \varphi(T)$$

for some known function φ, and this could be satisfied by taking, for example,

$$\sigma_T^2 = \varepsilon + \varphi(T)^-,$$
$$\alpha_T' + 2\beta_T \alpha_T = \varepsilon + \varphi(T)^+$$

for some $\varepsilon > 0$ fixed. This model is also clearly extremely *flexible*, in that any initial yield curve, and any initial term-structure of volatility can be fitted. However, the estimation of the model from data is not practical. Firstly, the yield curve is not some nice smooth curve known at all positive real points; in practice, it is only known at a limited set of maturities (typically 10-20), with dubious accuracy of measurement. Any procedure which requires repeated differentiation of this "curve" cannot be expected to work. Secondly, even if one could obtain estimates of the functions α, β, σ from the data, there is no reason why we should get consistent estimates if we performed the same analysis of the term-structure as it appears one week later!

The best we might hope for is to restrict (α, β, σ) to lie in some small parametric family, then estimate the parameters. The smallest interesting family we could consider is the Vasicek model, where the functions are constant, and the bond prices given by (2.7) simplify to

$$(2.11) \qquad m_t = \frac{\alpha}{\beta}t + \frac{1 - e^{-\beta t}}{\beta}\left(r_0 - \frac{\alpha}{\beta}\right)$$

$$(2.12) \qquad v_t = \frac{\sigma^2}{2\beta^3}\left[2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}\right].$$

This special case is concrete enough for us to investigate the only undesirable feature of these Gaussian models, namely that the interest rates may be *negative*. Now in the Vasicek model, the limiting distribution of r is $N\left(\alpha/\beta, \sigma^2/2\beta\right)$, so if we choose (α/β) to be reasonably large compared with the standard deviation $\sigma/\sqrt{2\beta}$, we might imagine that negative interest rates will not be a problem. However, taking

$$\alpha = \frac{1}{3} \times 10^{-3}, \quad \beta = \frac{1}{3} \times 10^{-2}, \quad \sigma^2 = \frac{8}{3} \times 10^{-6}$$

we have

(i) $\dfrac{\alpha}{\beta} = 5\dfrac{\sigma}{\sqrt{2\beta}}$ (ii) $\dfrac{\alpha}{\beta} = 0.1$ (iii) $\dfrac{1}{t}\log P(0,t) \to 0.02$.

The first tells us that in equilibrium the probability of a negative interest rate is very small (about 3×10^{-7} in fact), the second tells us that the mean value of r is 0.1, not an unreasonable annual rate – but the third tells us that the bond prices *grow* exponentially, which is absurd! Admittedly $P(0,t)$ will not climb back to 1 for quite a long time, but a model which can do this is a model which must either be rejected, or handled with caution. It is not enough simply to hope that the problem can be neglected.

We should consider the spot-rate process to be r_t^+ rather than r_t, but then the tractable Gaussian behaviour is gone. We can say the following, however.

PROPOSITION 2.1. *Suppose that r is a Gaussian process, $Er_t = \mu_t$, $\mathrm{cov}(r_s, r_t) = \rho_{st}$. Then*

$$
0 \;\leq\; E\exp\left(-\int_0^T r_u\,du\right) - E\exp\left(-\int_0^T r_u^+\,du\right)
$$

$$
\leq\; Ee^{-R_T}\left\{1 - \exp\left(-E\left(e^{-R_T}\int_0^T r_u^-\,du\right)\Big/Ee^{-R_T}\right)\right\}
$$

where $R_T \;\equiv\; \displaystyle\int_0^T r_u\,du;$

$$
(2.13) \qquad = \; Ee^{-R_T}\left\{1 - \exp\left(-\int_0^T \sqrt{\rho_{ss}}\,G\left(\frac{\mu_s - \int_0^T \rho_{st}\,dt}{\sqrt{\rho_{ss}}}\right)ds\right)\right\},
$$

where

$$
(2.14) \qquad G(a) \equiv E(W_1 - a)^+ = \frac{1}{\sqrt{2\pi}}e^{-a^2/2} - a\overline{\Phi}(a)
$$

with $\overline{\Phi}$ the tail of the standard normal distribution. Thus

$$
0 \;\leq\; 1 - E\left(\exp-\int_0^T r_u^+\,du\right)\Big/E\left(\exp-\int_0^T r_u\,du\right).
$$

$$
(2.15) \qquad \leq\; 1 - \exp\left[-\int_0^T \sqrt{\rho_{ss}}\,G\left(\frac{\mu_s - \int_0^T \rho_{st}\,dt}{\sqrt{\rho_{ss}}}\right)ds\right]
$$

REMARK 2.1. Using the inequality

$$
(2.16) \qquad G(a) \leq \frac{e^{-a^2/2}}{\sqrt{2\pi}(1 + a^2)} + a^-
$$

gives a good idea of the sizes of the quantities involved. For the example of the Vasicek model,

$$\mu_s \;=\; r_0 e^{-\beta s} + \frac{\alpha}{\beta}(1 - e^{-\beta s}),$$

$$\rho_{st} \;=\; \frac{\sigma^2}{2\beta}\left(e^{-\beta|s-t|} - e^{-\beta s - \beta t}\right),$$

so

$$\int_0^T \rho_{st}\,dt = \frac{\sigma^2}{\beta^2}\left\{1 - e^{-\beta s} - e^{-\beta T}\sinh\beta s\right\}.$$

Taking the earlier example with $\alpha = \frac{1}{3}\times 10^{-3}$, $\beta = 10\alpha$, $\sigma^2 = \frac{8}{3}\times 10^{-6}$, evaluating the bound for $T = 10$ gives 0.014, for $T = 20$ gives 0.041, for $T = 30$ gives 0.078 and for $T = 50$ gives 0.182.

Proof. The first statement is immediate, the second uses Jensen's inequality and for the third, we use the result that if X and Z are zero-mean Gaussians, $EX^2 = 1$, then for any θ

$$Ee^{-Z}(X + \theta)^- = \exp\left(\frac{1}{2}EZ^2\right)G(\theta - \operatorname{cov}(X, Z)).$$

$\square$

REMARK 2.2.   The estimate (2.13) is likely to be a good approximation to the difference between the bond prices using the Gaussian process r, and the positive part of r. This is because the only approximation used to reach (2.13) is $1 - e^{-x} \le x$. For small $x > 0$, the difference $x - 1 + e^{-x}$ is $O(x^2)$, and for larger values of x it is $O(x)$, but in the expectation, the distribution of r_s^- will have a rapidly decreasing tail if we have chosen parameters which make the probability of negative spot rates small. Thus the approximation should be accurate.

3. Squared Gaussian models. We can escape negative interest rates if we modify the variance structure in (2.1) to give

$$(3.1)\qquad\qquad dr_t = \sigma_t\sqrt{r_t}\,dW_t + (\alpha_t - \beta_r r_t)dt,$$

where once again $\alpha, \beta, \sigma : \mathbb{R}^+ \to \mathbb{R}^+$ are any locally bounded functions[1]. The process r will remain non-negative if it starts non-negative, a great advantage over the Gaussian models of the last section. Cox, Ingersoll & Ross [13] introduced such processes as models for the spot rate (taking α, β, σ to be constant); the time-dependent version which we take here is a generalisation due to Hull & White [27].

---

[1] It can be shown, using the Yamada-Watanabe theorem and time-change, that (3.1) has a pathwise unique strong solution if $\beta\sigma^{-2}$ is locally bounded; see, for example, V.26, V.40 of Rogers & Williams [38].

These models are particularly tractable because, like the Vasicek model, the yield curve is *affine in the spot rate*. More precisely, we have the following formula for the bond price:

$$(3.2) \qquad P(t,T) = \exp\{-r_t B(t,T) - A(t,T)\} \quad (0 \le t \le T),$$

where B solves the Riccati equation

$$(3.3) \quad \dot{B}(t,T) - \frac{1}{2}\sigma_t^2 B(t,T)^2 - \beta_t B(t,T) + 1 = 0, \quad B(T,T) = 0,$$

and A solves the simple first-order equation

$$(3.4) \qquad \dot{A}(t,T) = -\alpha_t B(t,T), \quad A(T,T) = 0.$$

It should be emphasized that the functions A, B used above are *not* the same as the functions A, B in the previous section ((2.9), (2.10)). Indeed, no simple closed form is available in general for A, B, though in the case of constant α, β, σ we have the (Cox-Ingersoll-Ross) formulae

$$(3.5) \qquad A(t,T) \;=\; -\frac{2\alpha}{\sigma^2} \log\left[\frac{\gamma e^{\beta\tau/2}}{\gamma \cosh \gamma\tau + \frac{1}{2}\beta \sinh \gamma\tau}\right],$$

$$(3.6) \qquad B(t,T) \;=\; \frac{\sinh \gamma\tau}{\gamma \cosh \gamma\tau + \frac{1}{2}\beta \sinh \gamma\tau},$$

where $\tau = T - t$, $2\gamma \equiv (\beta^2 + 2\sigma^2)^{1/2}$.

Compared to the simple Gaussian models of the last section, these models are *more realistic*, in that the spot rate stays non-negative, but *not so simple*, because one has to solve the Riccati equation (3.3). It is well known that (3.3) can be reduced to a second-order linear equation, and a brisk treatment of this is given in Appendix A.

All of the objections raised to the estimation of the Gaussian models of the last section apply equally well here, in particular, the functions α, β, σ cannot be uniquely determined from the term-structure and term-structure of volatility, and any attempt to recover them from the data involves differentiating up to three times. Jamshidian [28] has studied this class of models, and finds that if one restricts α, β, σ by insisting that α/σ^2 is constant, then certain simplifications result, and for given $A(0, \cdot), B(0, \cdot)$ there are unique α, β, σ satisfying $\alpha\sigma^{-2} = $ constant. To see why this restriction is natural, let us approach the problem from the other end, starting with the constant-coefficient Cox-Ingersoll-Ross model, and see what perturbations we could make to it. This work is done with Wolfgang Stummer.

One thing we could do is to make a *deterministic C^2 time-change*, and thus represent the bond prices as

$$(3.7) \qquad P(t,T) = E\left[\exp - \int_{\tau(t)}^{\tau(T)} r_u\, du \Big| \mathcal{F}_{\tau(t)}\right],$$

where r is a standard Cox-Ingersoll-Ross process, solving (3.1) with constant α, β, σ. If $P_0(x, \tau)$ denotes the price of a bond of maturity τ when the initial spot rate is x and the spot-rate process is r, then P_0 is available in closed form (from (3.2), (3.6)) and from (3.7) we have immediately that

$$(3.8) \qquad P(t, T) \quad = \quad P_0(r(\tau_t), \tau_T - \tau_t)$$

$$= \quad E\left[\exp - \int_t^T \tau_u' r(\tau_u) du | \mathcal{F}_{\tau(t)}\right]$$

$$(3.9) \qquad\qquad\quad = \quad P_0(\rho_t/\tau_t', \tau_T - \tau_t),$$

where $\rho_t \equiv \tau_t' r(\tau_t)$ is the spot rate process, relative to the given time scale. (To see this, observe that $-\frac{\partial}{\partial T} \log P(t, T)|_{T=t} \equiv f_{tt}$ is the spot rate, and now differentiate (3.8) with respect to T.)

Routine methods of stochastic calculus show that ρ solves a stochastic differential equation

$$(3.10) \qquad d\rho_t = \sigma \tau_t' \sqrt{\rho_t} dW_t + \left[\alpha(\tau_t')^2 - (\beta\tau_t' - \tau_t''/\tau_t')\rho_t\right] dt.$$

More important, however, is the fact that with this model, one may fit *any initial term structure exactly*. This is trivial; we know that for fixed t the function $T \mapsto P(t, T)$ decreases continuously to zero, and we use (3.8) and (3.9) to tell us what should be the function τ. (The initial slope τ_0' is indeterminate, and could be chosen to match up the volatility of the spot rate, for example.)

A second simple transformation of the basic CIR model that we could perform is to *multiply by some positive C^1 function g*. It is easy to check that if we define $\tilde{r}_t \equiv g_t \rho_t$, then $\tilde{r}$ solves the SDE (3.1) with

$$(3.11) \qquad\qquad \sigma_t \quad = \quad \sigma \tau_t' \sqrt{g_t},$$

$$(3.12) \qquad\qquad \alpha_t \quad = \quad \alpha \sigma_t^2 / \sigma^2,$$

$$(3.13) \qquad\qquad \beta_t \quad = \quad \beta \tau_t' - \frac{\tau_t''}{\tau_t'} - \frac{g_t'}{g_t}.$$

Given any functions $\sigma_\bullet > 0$ and $\beta_\bullet$, we can always choose τ increasing and $g > 0$ to make (3.11) hold.[2] To summarize then, *the processes satisfying* (3.1) *with* $\sigma_t^{-2}\alpha_t =$ constant *are exactly those obtained from a standard Cox-Ingersoll-Ross process by deterministic C^2 time-change, and multiplication by a deterministic C^1 function.*

REMARK 3.1.   Suppose that X is a Gauss-Markov process in $\mathbb{R}^d$ solving

$$dX_t = \frac{1}{2}\sigma_t dW_t - \frac{1}{2}\beta_t X_t dt.$$

---

[2] Write $v = \tau' > 0$, so that $v^2 g = \sigma_t^2/\sigma^2$ is known, and, multiplying (3.13) by vg we find $\beta_t v_t g_t = \beta(v_t^2 g_t) - (v_t g_t)' = \beta \sigma_t^2/\sigma^2 - (v_t g_t)'$; now this is easily solved for $v_t g_t$, and hence we deduce v_t, g_t.

Then if $r_t \equiv |X_t|^2$, and $\sigma, \beta : \mathbb{R}^+ \to \mathbb{R}$, simple Itô calculus reveals that

$$(3.14) \qquad dr_t = \sigma_t \sqrt{r_t} dZ_t + \left(\frac{d}{4} \sigma_t^2 - \beta_t r_t \right) dt$$

where W is Brownian motion in $\mathbb{R}^d$, Z is Brownian motion in $\mathbb{R}$. Comparing (3.14) with (3.1) shows that $4\alpha_t / \sigma_t^2$ is the "dimension" of the Gaussian process and so it becomes less surprising that the condition $\alpha\sigma^2 = $ constant should make an appearance. It also explains the title of this section! For those who are familiar with such things, we are in the land of Bessel processes (for a recent survey, see Revuz & Yor [36]), which also appear to be connected with the Ray-Knight theorems on Brownian local time (see, for example, Rogers & Williams [38]) and diffusion limits of branching processes; these things are all very closely related.

The squared-Gaussian and Gaussian models have an affine yield, but what other processes are there with this desirable property? This question has been answered in various forms (see Cox, Ingersoll & Ross [14], Brown & Schaefer [6] for example). Following [14], we shall consider the spot-rate process to solve the SDE

$$(3.15) \qquad dr_t = \sigma(t, r, \xi) dW_t + \mu(t, r, \xi) dt$$

where $\xi_t = (\xi_i(t))_{i=1}^n$ is a vector of exponentially-weighted past values of r,

$$(3.16) \qquad \xi_i(t) = e^{-\lambda_i t} \int_{-\infty}^{t} \lambda_i e^{\lambda_i u} r_u du$$

so that

$$(3.17) \qquad d\xi_i = \lambda_i (r - \xi_i) dt.$$

If now we require that

$$y(t, T) \equiv -\log P(t, T) = A(t, T) + B(t, T) r_t + \sum_{i=1}^{n} C_i(t, T) \xi_i(t)$$

then applying Itô's formula to the martingale $\exp\left(-\int_0^t r_u du \right) P(t, T)$ gives

$$(3.18) \quad \frac{1}{2} B(t, T)^2 \sigma(t, r, \xi)^2 \quad - \quad \dot{A}(t, T) - r - \dot{B}(t, T) r - \mu(t, r, \xi) B(t, T)$$

$$- \sum_{1}^{n} \{ \dot{C}_i(t, T) \xi_i + \lambda_i C_i(t, T)(r - \xi_i) \} = 0$$

where, as before, a dot denotes differentiation with respect to t. Hence

$$\mu(t, r, \xi) = \frac{1}{2} B(t, T) \sigma(t, r, \xi)^2 \quad - \quad \frac{\dot{A}}{B}(t, T) - \frac{r}{B(t, T)} - r \frac{\dot{B}}{B}(t, T)$$
$$- \quad \sum_1^n \left\{ \frac{\dot{C_i}}{B}(t, T) \xi_i + \lambda_i \frac{C_i}{B}(t, T)(r - \xi_i) \right\}.$$

Now the right-hand side of this is the same for all $T > t$, so taking the difference for two values of $T > t$ we deduce that $\sigma(t, r, \xi)^2$ must be of the form

$$(3.19) \qquad \sigma(t, r, \xi) = \sigma_0(t)^2 + \sigma(t)^2 r + \sum_1^n \sigma_i(t)^2 \xi_i,$$

and from this

$$(3.20) \qquad \mu(t, r, \xi) = \mu_0(t) + \mu(t) r + \sum_1^n \mu_i(t) \xi_i.$$

Returning these to (3.18) yields differential equations for A,B,C, which can always be solved numerically.

REMARK 3.2.   (i) The Gaussian models appear if we take $\sigma_i = 0$, $\sigma = 0$ in (3.19), and $\mu_i = 0$ in (3.20). The squared-Gaussian models arise when we take $\sigma_0 = 0 = \sigma_i$ in (3.19).

(ii) If we were to insist that $\sigma_0(\cdot)$, $\sigma(\cdot)$, $\sigma_i(\cdot)$, $\mu_0(\cdot)$, $\mu(\cdot)$, and $\mu_i(\cdot)$ were all constants, we have a wide range of possible models available to fit any given yield curve, though the analysis will not be particularly simple. Nonetheless, it may be preferable to use one of this class of models rather than one of the time-dependent versions.

(iii) All of these models are *single-factor* models, and can be criticized on these grounds. Other single-factor models have been proposed by Dothan [16], Brennan & Schwartz [4], Courtadon [11], and Cox [12] (see also Beckers [3]); none of these appears to be conclusively superior to the models discussed already, and all have lost the analytical tractability of the models discussed above.

4. Multi-factor models. The single-factor Markovian models discussed so far are often criticized on the grounds that the long rate is a deterministic function of the spot rate, and that the prices of bonds of different maturities are perfectly correlated. These defects might perhaps be forgiven if the models appeared to match observed prices, but, as we shall review in Section 7 the weight of empirical evidence suggests that multi-factor models do significantly better than single-factor models.

In the last few years, there has been a rash of papers all dealing with the same broad class of models, the higher-dimensional squared-Gauss-Markov

processes. This title is longer (and less informative) than the SDE defining them:

$$(4.1) \qquad dX_t \;\;=\;\; \sigma_t dW_t + (a_t + C_t X_t) dt$$

$$(4.2) \qquad r_t \;\;=\;\; \frac{1}{2}|X_t|^2,$$

where W is a Brownian motion in $\mathbb{R}^n$, and $\sigma, C : \mathbb{R}^+ \to \mathbb{R}^n \otimes \mathbb{R}^n$, $a : \mathbb{R}^1 \to \mathbb{R}^n$ are deterministic locally bounded functions. As was remarked above, the Cox-Ingersoll-Ross single factor models are obtained in this way if their dimension is integer. The papers known to me which deal with this class of models are Beaglehole & Tenney [2], El Karoui, Myneni & Viswanathan [21], Duffie & Kan [17], Constantinides [10], Jamshidian [28], although the starting points do differ in these various papers. They all end up with essentially the same class of models, examples of which were already studied by Cox, Ingersoll & Ross [13], Richard [37], and Longstaff & Schwartz [33].

The attraction of this class is that there is a (semi-)explicit formula for the bond price: it is easy to prove that

$$(4.3) \qquad P(t,T) = \exp\left\{ -\frac{1}{2} X_t^T Q_t X_t + b_t^T X_t - \gamma_t \right\},$$

where Q solves the matrix Riccati equation

$$(4.4) \qquad I + QC + (QC)^T + \dot{Q} - Q\sigma\sigma^T Q^T = 0, \quad Q(T,T) = 0,$$

and b, γ solve

$$(4.5) \qquad \dot{b} \;\;-\;\; Qa - (Q\sigma\sigma^T - C^T)b = 0, \quad b(T,T) = 0,$$

$$(4.6) \qquad \dot{\gamma} \;\;=\;\; b^T a - \frac{1}{2}tr(\sigma^T Q\sigma) + \frac{1}{2}b^T \sigma\sigma^T b, \quad \gamma(T,T) = 0.$$

This class of processes, and the differential equations (4.4)-(4.6), have been around for a long time in the world of stochastic processes, going back at least to Cameron & Martin [7]; see also Feller [22], Liptser & Shiryaev [31], and, for recent work, Donati-Martin & Yor [15], Rogers & Shi [39], Chan, Dean, Jansons & Rogers [9].

Though the matrix Riccati equation (4.4) can easily be solved numerically, there are only closed-form analytic solutions when the problem is one-dimensional, or can be reduced to independent one-dimensional processes; in general, the analysis of this class of models is sticky. Interest in these models will undoubtedly continue for some time to come; Duffie & Ken [17] have formulated the most general affine yield multifactor model, which appears to have closed-form solutions only in the special case of independent CIR processes (or some equivalent situation). Nevertheless, numerical solution is always a possibility.

By passing to multi-factor models, one should get an improved fit, but there is a heavy price to pay; if one wants to calculate prices of, say, options on bonds, the PDE to be solved is higher-dimensional, and will thus be much slower. Perhaps even more importantly, the factors used have to correspond to some observable variables if the formulae are ever to be used. Cox, Ingersoll & Ross [13] and Richard [37] both take the spot rate together with the rate of inflation, Longstaff & Schwartz [33] use the spot rate together with the volatility of the spot rate, and Duffie & Kan [17] use the yields on a fixed set of bonds, for example.

Outside of this class of squared Gaussian models, there are many which have been proposed, but I highlight just two. The first of these is the model of Brennan & Schwartz [4], which takes as the variables of the (two-factor) model the spot rate r and the long rate. As a proxy for the long rate, Brennan & Schwartz use the reciprocal of the price of a consol, which is a traded asset. This is a nice idea, and clarifies the analysis, though their pricing equation still has to be solved numerically. They use the model to analyze Canadian Government bonds, and obtain impressive results. A variant of the basic model is subsequently studied by Schaefer & Schwartz [40], who take the spot rate and the spread as the variables. In this context, it is worth recalling a result of Dybvig, Ingersoll & Ross [19], who prove that the long rate is non-decreasing. This makes one a little wary about a model which supposes that the long rate moves as a diffusion, even if only in the form of its proxy, the reciprocal of the consol.

The final model is the model of Fong & Vasicek [23], who take the spot rate r to follow the Ornstein-Uhlenbeck SDE (2.1) with constant α, β, but with σ itself following an independent squared-Gaussian diffusion. The model has a simple affine yield structure, which is attractive, but also the risk of negative spot rates. Whether this model is any better than a two-factor squared Gaussian model cannot be decided at a theoretical level, and must be resolved by comparing results on real data.

5. **Whole-yield models.** The approach adopted in whole-yield models is to model directly the forward-rate processes $(f_{tT})_{0 \leq t \leq T}$ for each T. The earliest (discrete-time) appearance of this approach was due to Ho & Lee [26]. In view of the fact (proved by Dybvig [18] and Jamshidian [28]) that the continuous-time limit of the Ho & Lee model is

$$dr_t = \theta_t dt + \sigma dW_t$$

for some deterministic function θ, it has not been adopted unreservedly. This possibility of spectacularly negative spot rates, together with the single-factor nature of the model, has led to more refined models, the main one being the model of Heath, Jarrow & Morton [25]. The same model was developed independently and contemporaneously by Babbs [1]. The idea

here is to model the forward-rate curve by

$$(5.1) \qquad f_{tT} = f_{0T} + \int_0^t \sigma(s, T)dW_s + \int_0^t \alpha(s, T)ds \quad (0 \le t \le T),$$

where W is an n-dimensional Brownian motion, and $(f_{0T})_{T \ge 0}$ is the initial forward-rate curve. The functions σ and α cannot be chosen unrestrictedly; indeed, Heath, Jarrow & Morton prove that in fact

$$(5.2) \qquad \alpha(t, T) = \sigma(t, T) \left\{ \varphi_t + \int_t^T \sigma(t, s)ds \right\},$$

where φ is some n-vector previsible process which is zero when working in the risk-neutral measure. In fact, this structure becomes almost obvious when looked at in the right way. Indeed, if we consider the martingale (with respect to the risk-neutral measure)

$$(5.3) \; M_t \equiv E \left[\exp \left(- \int_0^T r_u \, du \right) | \mathcal{F}_t \right] = \exp \left\{ - \int_0^t r_u \, du - \int_t^T f_{tu} \, du \right\}$$

we know that it can be represented in the form

$$(5.4) \qquad M_t = M_0 \exp \left\{ - \int_0^t \Sigma(s, T)dW_s - \frac{1}{2} \int_0^t |\Sigma(s, T)|^2 ds \right\}$$

for some previsible n-vector process $\Sigma(\cdot, T)$ (any martingale on the Brownian filtration can be represented as a stochastic integral, and hence easily any non-negative martingale on the Brownian filtration is representable as an exponential of a stochastic integral.) Taking the two expressions of $\log M_t$ from (5.3) and (5.4) and comparing the martingale parts shows that

$$\Sigma(t, T) = \int_t^T \sigma(t, u)du,$$

and now differentiating with respect to T gives (5.2), and also (5.1) when we remember that $\log M_0 = \int_0^t f_{0u} du$.

Of course, various regularity conditions are needed to justify this, and indeed further regularity conditions (which are not made explicit in [25]) are need to ensure that

$$\lim_{T \downarrow t} f_{tT} = r_t$$

exists, is continuous and defines a semimartingale (none of which need hold in general). However, such points are trivial in comparison to real objections to the use of the model in practice. Heath, Jarrow & Morton give two examples where there are simple formulae for bond prices, but where r is allowed to go negative; and they give an example where $r \geq 0$, but there are no simple formulae. It appears very difficult to obtain both of the desirable properties together, not least because this approach begins by trying model *derived* quantities (the forward rates) instead of the *fundamental* quantity (the spot rate) and thus loses control.

If one ever *could* find a specification of the model where $r \geq 0$ and where bond prices were given by a simple formula, then one could just as well obtain this by starting directly with the spot rate process r! So these whole-yield models appear to offer no advantage over the approach of modelling r, although the ability to input the initial yield curve directly and to vary the volatilities of different forward rates are attractions (partly shared by some of the models of the preceding sections.)

6. Equilibrium or arbitrage pricing? Cox, Ingersoll & Ross [13] in their paper on term structure of interest rates discussed (in Section 5) the comparison of their own equilibrium approach and the arbitrage approach. What they wrote there appears to have caused considerable confusion about the relation between the two; the aim of this section is to prove that the two are essentially equivalent.

To begin with, we suppose that $(S_t)_{0 \leq t \leq T} = ((S_t^0, \ldots, S_t^n))_{0 \leq t \leq T}$ is an $(n+1)$-vector continuous semimartingale of *financial asset* prices. The first component S^0 is non-decreasing, and represents the price of a riskless bond. There will be a single *productive asset* whose price at time t is ξ_t, where ξ is also a continuous semimartingale. Prices are in terms of the unique commodity of this economy. An investor starts with wealth x, and holds a self-financing portfolio (θ_t, H_t) in the financial assets and the productive asset respectively. He consumes at rate $C_t \geq 0$ at time t, so his wealth at time t, X_t, obeys the wealth equation

$$(6.1) \qquad dX_t = \theta_t dS_t + H_t d\xi_t - C_t dt,$$

$$(6.2) \qquad X_t = \theta_t S_t + H_t \xi_t, \quad X_0 = x.$$

The investor aims to choose (θ, H) so as to

$$(6.3) \qquad \max E \int_0^T u(C_s) ds \text{ subject to } X_t \geq 0 \text{ for all } t \in [0, T].$$

We shall say that the price processes (S, ξ) constitute an *equilibrium* for this model if under optimal play the investor invests nothing in the financial assets. This terminology bears the usual interpretation; we imagine a number of identical investors, all investing in the market. The financial

assets only exist by virtue of some agents going short, some going long; the net supply of financial assets is zero. The productive asset has a physical existence, however.

We assume the utility function in (6.3) to be strictly increasing, C^2, strictly concave and unbounded above. Thus if (S, ξ) is an equilibrium, there must be no arbitrage (otherwise the investor would be able to make unlimited gain, and his utility-maximization problem would be ill-posed). It is a folk theorem of the subject that no arbitrage implies the existence of an equivalent martingale measure. Without pausing to examine the exact result (which, in any case, has not yet been formulated correctly), we concentrate on the converse.

THEOREM 6.1. *Write $\beta_t \equiv 1/S_t^0$, and suppose that there exists a measure $\tilde{P}$ equivalent to P such that under $\tilde{P}$ the process $\tilde{S} \equiv \beta S$ is a local martingale. Define the P-martingale Z by*

$$Z_t = \frac{d\tilde{P}}{dP}\bigg|_{\mathcal{F}_t}.$$

If we now take $U(x) \equiv \log x$, $\xi_t \equiv 1/\beta_t Z_t$, then the optimal policy for the investor is

$$(6.4) \qquad C_t^* = \frac{x}{T}\xi_t, \quad \theta_t^* = 0, \quad H_t^* = x\left(1 - \frac{t}{T}\right).$$

REMARK 6.1. What this says is that if there exists an equivalent measure under which all the financial asset prices, when discounted, are local martingales, *then there is an economy which supports these as equilibrium prices*: The reason that the example of [13] does not violate this is as follows. Cox, Ingersoll & Ross are working in a Brownian framework, where a change of measure corresponds to introducing a drift into the Brownian motion. However, not *every* possible drift corresponds to a change of measure, and what they have done is to consider a drift which does not correspond to such a change of measure! To quote Cox, Ingersoll & Ross, "The difficulty, of course, is that there is no underlying equilibrium which would support the assumed premiums" – but, on the other hand, there is no risk-neutral measure either!

Proof. With the portfolio and consumption plan (6.4), the wealth process is

$$X_t = x(1 - t/T)\xi_t$$

and it is two lines of calculus to confirm that this solves the wealth equation (6.1). Reworking the wealth equation (6.1) for general portfolio/consumption,

$$(6.5) \qquad d(\beta_t X_t) = \theta_t d\tilde{S}_t + H_t d\tilde{\xi}_t - \beta_t C_t dt$$

(where $\tilde{\xi}_t \equiv \beta_t \xi_t \equiv 1/Z_t$) so that

$$0 \leq \beta_t X_t = x + \int_0^t (\theta_u d\tilde{S}_u + H_u d\tilde{\xi}_u) - \int_0^t \beta_u C_u du.$$

Now under $\tilde{P}$, both $\tilde{S}$ and $\tilde{\xi}$ are local martingales, so by Fatou's lemma

$$(6.6) \qquad x \geq \tilde{E}\left(\int_0^T \beta_u C_u du\right) = E\left(\int_0^T \beta_u Z_u C_u du\right)$$

for any feasible consumption plan. The proof is completed by a simple "Lagrangian sufficiency" argument. As a piece of notation, we set for $x > 0$

$$(6.7) \qquad U_*(x) \quad = \quad \sup\{U(c) - xc\}$$
$$(6.8) \qquad \qquad \quad = \quad U(I(x)) - xI(x),$$

where I is the inverse to U' (in fact, the second line only holds if U' decreases from infinity at 0+ to zero at infinity, which is certainly true when, as we assume here, $U = \log)^3$. Now for any $\lambda > 0$, we have for any feasible consumption process

$$(6.9)\, E\int_0^T U(C_s)ds \quad \leq \quad E\left[\int_0^T U(C_s)ds + \lambda\left(x - \int_0^T \beta_u Z_u C_u du\right)\right]$$
$$= \quad \lambda x + E\int_0^T \{U(C_s) - \lambda\beta_s Z_s C_s\}ds$$
$$\leq \quad \lambda x + E\int_0^T U_*(\lambda\beta_s Z_s)ds$$
$$= \quad \lambda x + E\int_0^T \{U(I(\lambda\beta_s Z_s)) - \lambda\beta_s Z_s I(\lambda\beta_s Z_s)\}ds.$$

Now if we write $C_s^\lambda \equiv I(\lambda\beta_s Z_s)$, then provided λ were chosen so that the "budget constraint" (6.6) is satisfied with equality, that is

$$(6.10) \qquad x = E\int_0^T \beta_s Z_s C_s^\lambda ds \equiv E\int_0^T \beta_s Z_s I(\lambda\beta_s Z_s)ds,$$

---

[3] U_* is the convex dual of U, and $U(x) = \inf_y\{U_*(y) - xy\}$

then (6.9) becomes simply

$$E \int_0^T U(C_s^\lambda) ds.$$

However, we *can* achieve (6.10) in general, and trivially in this case where $I(x) = 1/x$ simply by choosing $\lambda = T/x$! Assembling the inequality, we have to conclude that for any feasible consumption plan C

$$E \int_0^T U(C_s) ds \leq E \int_0^T U(C_s^{T/x}) ds \equiv E \int_0^T U(C_s^*) ds,$$

and the proof is complete. □

Dybvig & Ross [20] have already remarked on the essential equivalence of 'equilibrium' and 'arbitrage' pricing. I understand that Heston has a similar argument, reported in Exercise 9.3 of Duffie's book *Dynamic Pricing Theory*.

7. Empirical aspects. A recent search of a computer data base turned up 135 articles which referenced the fundamental paper [13] of Cox, Ingersoll & Ross; of these, about 75% were principally theoretical. Of the papers which are principally empirical, I discuss here only a few, but I think the conclusions reported give a good feel for what has been deduced from data so far; no strongly-preferred model emerges, but some models appear to be inadequate.

The first problem to be faced is that there are very few zero-coupon bonds traded; US Treasury bills seem to be the only ones for which much data is available, and the maturities only go out to about a year. One is then faced with the problem of estimating the yield curve from other (coupon-bearing) bonds, and possibly other information. Anyone who works on interest rates, be they practitioner or academic, has a way of making a yield curve from such data, or has a source of yield-curve data where this information has already been stripped out or "estimated" and we shall not discuss further how these are obtained.

Most of the papers have some model, or class of models, in mind, and proceed to test the model, or some feature of it. The most common class of models is the class of squared-Gaussian models with one or more factors, though Chan, Karolyi, Longstaff & Sanders [8] cast their net wider and take their class to be (single-factor) models where r solves an SDE of the form

$$dr_t = \sigma r_t^\gamma dW_t + (\alpha + \beta r_t) dt$$

where the parameters $\alpha, \beta, \gamma > \frac{1}{2}$, σ are to be estimated. This includes the CEV model of Cox [12] (see Beckers [3] for a description of the model),

for example. The one-factor squared-Gaussian (Cox-Ingersoll-Ross) model is tested by Brown & Dybvig [5] using a least-squares fit, and by Pearson & Sun [35] using exact maximum likelihood. Both conclude that the one-factor model does not satisfactorily fit the data. On the other hand, Gibbons & Ramaswamy [24] find that it does perform quite well on the short-term Treasury bill data. Multifactor squared Gaussian models are tested by Stambaugh [41], Longstaff & Schwartz [33], and Litterman, Scheinkman & Weiss [32], who all conclude that introducing additional factors significantly improves the fit; in fact, using two factors appears to be satisfactory from the work of these quthors, although Pearson & Sun [35] do not find that two factors are sufficient.

The generalized method of moments is a popular approach to the estimation. Despite the arbitrariness of the procedure, it has some attractive features; the large-sample behaviour does not depend on specific distributional assumptions, needing only that the spot rate process be stationary and ergodic. A little care is needed here; Chan et at. [14] conclude that Dothan's [16] model, the Cox CEV [13] and the Cox-Ingersoll-Ross variable-rate model all do better than Vasicek or the standard Cox-Ingersoll-Ross squared-Gaussian, but these three diffusions are not ergodic, so the results must be interpreted with caution.

The outcome of the empirical studies seems to be that a two- or three-factor squared-Gaussian model is reasonably satisfactory, but there is one conclusion that most agree on, namely, that more work is needed here!

8. Conclusions: where now. Long before stochastic calculus hit the industry, bonds were being traded; practitioners had a good idea what prices to charge, and were exploiting their knowledge of the market, and hunches about the future, to guide them. It is futile to imagine that increasingly sophisticated mathematical models will replace or displace such skill, and what is needed now is not *more* (and more complicated) mathematical models, but rather a serious attempt to combine practitioner input with (probably extremely simple) mathematical models; no mathematical model based on assumptions of time-homogeneity can ever represent a world of elections, trade figures, summits and treaties.

One thing which does appear to be well worth doing (of a more academic nature) is to try to model index-linked bonds. Asking practitioners what they consider the main influence on term-structure, the response is that anticipated inflation is the most important effect. By removing that, through studying index-linked bonds, we may be able to see a more orderly pattern emerging, just as one does with share prices when the "operational time" effect is removed (see, for example, the working paper "Some statistics for testing the influence of the number of transactions on the distributions of returns" by S.E. Satchell & Y. Yoon.).

REFERENCES

[1] S. BABBS, *A family of Itô process models for the term structure of interest rates*, University of Warwick, preprint 90/24, 1990.

[2] D.R. BEAGLEHOLE & M.S. TENNEY, *General solution of some interest rate-contingent claim pricing equations*, J. Fixed Income **1**, 1991, pp. 69–83.

[3] S. BECKERS, *The constant elasticity of variance model and its implications for option pricing*, J. Finance **35**, 1980, pp. 661–673.

[4] M.J. BRENNAN & E.S. SCHWARTZ, *A continuous time approach to the pricing of bonds*, J. Banking Fin. **3**, 1979, pp. 133–155.

[5] S.J. BROWN & P.H. DYBVIG, *The empirical implications of the Cox-Ingersoll-Ross theory of the term structure of interest rates*, J. Finance, **41**, 1986, pp. 617–630.

[6] R.H. BROWN & S.M. SCHAEFER, *Interest rate volatility and the term structure*, London Business School preprint, 1991.

[7] R.H. CAMERON & W.T. MARTIN, *Evaluation of various Wiener integrals by use of certain Sturm-Liouville differential equations*, Bull. Amer. Math. Soc. **51**, 1945, pp. 73–90.

[8] K.C. CHAN, G.A. KAROLYI, F.A. LONGSTAFF, & A.B. SANDERS, *An empirical comparison of alternative models of the short-term interest rate*, J. Finance **47**, 1992, pp. 1209–1227.

[9] T. CHAN, D. DEAN, K.M. JANSONS, & L.C.G. ROGERS, *On polymer conformations in elongational flows*, Comm. Math. Phys. **160**, 1994, pp. 239–257.

[10] G.M. CONSTANTINIDES, *A theory of the nominal term structure of interest rates*, Rev. Fin. Studies **5**, 1992, pp. 531–552.

[11] G. COURTADON, *The pricing of options on default-free bonds*, J. Fin. Quantit. Anal. **17**, 1982, pp. 75–100.

[12] J.C. COX, *Notes on option pricing I: constant elasticity of variance diffusions*, Stanford University working paper, 1975.

[13] J.C. COX, J.E. INGERSOLL, & S.A. ROSS, *A theory of the term structure of interest rates*, Econometrica **53**, 1985, pp. 385–408.

[14] J.C. COX, J.E. INGERSOLL, & S.A. ROSS, *A re-examination of traditional hypotheses about the term structure of interest rates*, J. Finance **36**, 1981, pp. 769–798.

[15] C. DONATI-MARTIN & M. YOR, *Fubini's theorem for double Wiener integrals and the variance of the Brownian path*, Ann. Inst. Henri Poincaré **27**, 1991, pp. 181–200.

[16] L.U. DOTHAN, *On the term structure of interest rates*, J. Fin. Econ. **6**, 1978, pp. 59–69.

[17] D. DUFFIE & R. KAN, *A yield factor model of interest rates*, GSB Stanford University preprint, 1993.

[18] P.H. DYBVIG, *Bond and bond option pricing based on the current term structure*, Washington University working paper, 1988.

[19] P.H. DYBVIG, J.E. INGERSOLL, & S.A. ROSS, *Long forward rates can never fall*, Washington University working paper, 1992.

[20] P.H. DYBVIG & S.A. ROSS, *Arbitrage: The new Palgrave: a Dictionary of Economics*, Stockton Press, New York, 1987, pp. 100–106.

[21] N. EL KAROUI, R. MYNENI, & R. VISWANATHAN, *Arbitrage pricing and hedging of interest claims with state variables I: Theory*, Preprint, 1992.

[22] FELLER, W., *Two singular diffusion problems*, Ann. Math. **54**, 1951, pp. 173–182.

[23] H.G. FONG & O.A. VASICEK, *Interest-rate volatility as a stochastic factor*, Gifford Fong Associates working paper, 1991.

[24] M.R. GIBBONS & K. RAMASWAMY, *The term structure of interest rates: empirical evidence*, Stanford University working paper, 1986.

[25] D. HEATH, R. JARROW, & A. MORTON, *Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation*, Econometrica

60, 1992, pp. 77–105.

[26] T.S.Y. HO & S.-B. LEE, *Term structure movements and pricing interest rate contingent claims*, J. Finance **41**, 1986, pp. 1011–1029.

[27] J. HULL & A. WHITE, *Pricing interest-rate derivative securities*, Rev. Fin. Studies, **3** 1990, pp. 573–592.

[28] F. JAMSHIDIAN, *The one-factor Gaussian interest rate model: Theory and implementation*, Working paper, Merrill Lynch Capital Markets, 1988.

[29] F. JAMSHIDIAN, *A simple class of square-root interest rate models*, Working paper, Fuji International Finance, 1993.

[30] F. JAMSHIDIAN, *Bond, futures and option evaluation in the quadratic interest rate model*, Working paper, Fuji International Finance, 1993.

[31] R.S. LIPTSER & A.N. SHIRYAEV, *Statistics of Random Processes I*, Springer, Berlin, 1977.

[32] R. LITTERMAN, J. SCHEINKMAN, & L. WEISS, *Volatility and the yield curve*, Working paper, Goldman Sachs & Co., New York, 1988.

[33] F.A. LONGSTAFF & E.S. SCHWARTZ, *Interest-rate volatility and the term structure: a two factor general equilibrium model*, J. Finance **47**, 1991, pp. 1259–1282.

[34] R.C. MERTON, *Theory of rational option pricing*, Bell J. Econ. Man. Sci. **4**, 1973, pp. 141–183.

[35] N.D. PEARSON & T.-S. SUN, *A test of the Cox-Ingersoll-Ross model of the term structure of interest-rates using the method of maximum likelihood*, Working paper, MIT Sloan School, 1989.

[36] D. REVUZ & M. YOR, *Continuous Martingales and Brownian Motion*, Springer, 1991.

[37] S.F. RICHARD, *An arbitrage model of the term structure of interest rates*, J. Fin. Econ. **6**, 1978, pp. 33–57.

[38] L.C.G. ROGERS & D. WILLIAMS, *Diffusion Markov Processes and Martingales, Vol. 2*, Wiley, Chichester, 1987.

[39] L.C.G. ROGERS & Z. SHI, *Quadratic functionals of Brownian motion, optimal control, and the "Colditz" example*, Stochastics and Stochastics Reports **41**, 1992, pp. 201–218.

[40] S.M. SCHAEFER & E.S. SCHWARTZ, *Time-dependent variance and the pricing of bond options*, J. Finance **42**, 1987, pp. 1113–1128.

[41] R.F. STAMBAUGH, *The information in forward rates: implications for models of the term structure*, J. Fin. Econ. **21**, 1988, pp. 41–70.

[42] O.A. VASICEK, *An equilibrium characterization of the term structure*, J. Fin. Econ. **5**, 1977, pp. 177–188.

A. Appendix. (i) We take r to solve the SDE

$$(A.1) \qquad dr_t = \sigma_t \sqrt{r_t} dW_t + (\alpha_t - \beta_t r_t) dt$$

so changing variables to $z_t \equiv 2r_t / \sigma_t^2$, we find that

$$(A.2) \qquad dz_t = \sqrt{2z_t} dW_t + \left[\frac{2\alpha_t}{\sigma_t^2} - \beta_t z_t - 2\dot{\sigma}_t \sigma_t^{-1} z_t \right] dt.$$

For notational simplicity, set $a_t \equiv 2\alpha_t \sigma_t^{-2}$, $b_t \equiv \beta_t + 2\dot{\sigma}_t \sigma_t^{-1}$, so that the bond price $P(z, t, T)$ ($0 \leq t \leq T$) satisfies for fixed T

$$(A.3) \qquad \dot{P} + zP'' + (a - bz)P' - \frac{1}{2}\sigma^2 zP = 0,$$

where a dot denotes differentiation with respect to t, and a dash denotes differentiation with respect to z. This PDE has a solution of the form

$$P(z, t, T) = \exp\left[-z_t B(t, T) - A(t, T) \right]$$

provided

$$(A.4) \qquad \dot{B} - B^2 - bB + \frac{1}{2}\sigma^2 = 0, \qquad B(T, T) = 0$$

$$(A.5) \qquad \dot{A} = -aB, \quad A(T, T) = 0.$$

(Note that, because of the change of variables to z, what was denoted by $B(t, T)$ in Section 3 is here denoted $2B(t, T)\sigma_t^{-2}$; the change of notation should cause no confusion, and is particularly convenient for the purposes of this section.)

Writing

$$(A.6) \qquad B(t, T) = -\dot{\psi}(t, T)/\psi(t, T),$$

we can recast (A.4) as

$$(A.7) \qquad \ddot{\psi} - b\dot{\psi} - \frac{1}{2}\sigma^2 \psi = 0.$$

Let us write $\psi_{\pm}$ for two linearly independent solutions of this second-order linear differential equations; we shall choose ψ_+ to be increasing non-negative, ψ_- to be decreasing non-negative, $\psi_+(0) = \psi_-(0) = 1$, $\psi_+'(0) = 0$.[4] In terms of these, if we set

$$(A.8) \qquad \psi(t, T) = \psi_-(t)\dot{\psi}_+(T) - \psi_+(t)\dot{\psi}_-(T)$$

---

[4] To see that such a choice is possible, if we let x be the solution to the SDE

$$dx_t = dW_t - b(x_t, T)\,\mathrm{sgn}(x_t)dt$$

and if $H_a \equiv \inf\{t : x_t = a\}$, $\varphi_t = \int_0^t \frac{1}{2}\sigma(x_u, T)^2 du$, then we can take $\psi_-(x) = E_x \exp\{-\varphi(H_0)\}$, $\psi_+(x) = 1/E^0 \exp\{-\varphi(H_x \wedge H_{-x})\}$.

then

$$(A.9) \qquad B(t,T) = -\frac{\dot\psi(t,T)}{\psi(t,T)} = -\frac{\partial}{\partial t}\log\psi(t,T)$$

is the solution to the Riccati equation, and the form of A is

$$(A.10) \qquad A(t,T) = \int_t^T a_u B(u,T)du.$$

(ii) It is of interest to consider the inverse problem, that is, if we are told the initial functions $B(0,\cdot)$, $A(0,\cdot)$, can we find some process of the form (A.1) which would give these A, B? This needs an understanding of how $B(t,T)$ varies with T. Now

$$
\begin{aligned}
\frac{\partial}{\partial T}B(t,T) &\equiv B_T(t,T) \\
&= -\frac{\partial^2}{\partial t\,\partial T}\log\psi(t,T) \\
&= -\frac{\partial}{\partial t}\frac{\psi_-(t)\ddot\psi_+(T) - \psi_+(t)\ddot\psi_-(T)}{\psi_-(t)\dot\psi_+(T) - \psi_+(t)\dot\psi_-(T)} \\
&= -\frac{\partial}{\partial t}\left[b_T + \frac{1}{2}\sigma_T^2 \frac{\psi_-(t)\psi_+(T) - \psi_+(t)\psi_-(T)}{\psi_-(t)\dot\psi_+(T) - \psi_+(t)\dot\psi_-(T)}\right] \\
(A.11) \qquad &= \frac{1}{2}\sigma_T^2 \frac{\xi_t\xi_T}{\psi(t,T)^2},
\end{aligned}
$$

where

$$(A.12) \qquad \xi_t = \psi_-(t)\dot\psi_+(t) - \psi_+(t)\dot\psi_-(t) = \psi(t,t).$$

If we define $\gamma(t,T) \equiv \psi_-(t)\psi_+(T) - \psi_+(t)\psi_-(T)$, then we have after some calculations

$$
\begin{aligned}
(A.13) \qquad \frac{\partial}{\partial T}\log B_T(t,T) &= \frac{2\dot\sigma_T}{\sigma_T} - b_T - \sigma_T^2\frac{\gamma(t,T)}{\psi(t,T)} \\
&= -\beta_T - \sigma_T^2\frac{\gamma(t,T)}{\psi(t,T)}
\end{aligned}
$$

and differentiating once more, and rearranging, one gets after some calculations

$$(A.14) \quad \frac{\partial^2}{\partial T^2}\log B_T(t,T) = -\beta_T' - \sigma_T^2 - \frac{1}{2}\beta_T^2 + \frac{1}{2}\left(\frac{\partial}{\partial T}\log B_T(t,T)\right)^2.$$

If we were in the special case where a was constant, then knowing $A(0,\cdot)$ would tell us, from (A.9), (A.10), the function

$$\log\psi(T,T) - \log\psi(0,T)$$

and if we differentiate this with respect to T, we get

$$-\frac{1}{2}\sigma_T^2 \frac{\gamma(0,T)}{\psi(0,T)}.$$

Combining with (A.13) would tell us what the function β should be, and now returning this to (A.14), we deduce the function σ^2. But notice that to recover the coefficient functions σ^2, β, we have had to differentiate the bond prices three times.