A Yield-Factor Model of Interest Rates

Darrell Duffie and Rui Kan

Graduate School of Business, Stanford University

June, 1992; revised: August, 1995

Abstract: This paper presents a consistent and arbitrage-free multi-factor model of the term structure of interest rates in which yields at selected fixed maturities follow a parametric multi-variate Markov diffusion process, with "stochastic volatility." The yield of any zero-coupon bond is taken to be a maturity-dependent affine combination of the selected "basis" set of yields. We provide necessary and sufficient conditions on the stochastic model for this affine representation. We include numerical techniques for solving the model, as well as numerical techniques for calculating the prices of term-structure derivative prices. The case of jump-diffusions is also considered.

Please address all correspondence to: Darrell Duffie, Graduate School of Business, Stanford University, Stanford CA 94305-5015. We are grateful for discussions with Ken Singleton, Bob Litterman, Antoine Conze, Nicole El Karoui, Vincent Lacoste, Jeremy Evnine, Antoine Frachot, Henri Pagès, Jean-Philippe Lesne, Fischer Black, Ayman Hindy, George Pennachi, Rob Bliss, Prasad Nannisetty, Stan Pliska, and Chris Rogers, and especially to a referee for pointing out an error corrected in this version.

1. Introduction

This paper defines and analyses a simple multi-factor model of the term structure of interest rates. The factors of the model are the yields $X = (X_1, X_2, \ldots, X_n)$ of zero-coupon bonds of n various fixed maturities, $\{\tau_1, \tau_2, \ldots, \tau_n\}$. For example, one could think of the current 5-year (zero-coupon) yield as a factor. The yield factors form a Markov process, to be described below, that can be thought of as a multivariate version of the single-factor model of Cox, Ingersoll, and Ross (1985a). As opposed to most multi-factor term structure models, the factors (Markov state variables) are observable from the current yield curve and their increments can have an arbitrarily specified correlation matrix. The model includes stochastic volatility factors that are specified linear combinations of yield factors. Discount bond prices at any maturity are given as solutions to Ricatti (ordinary differential) equations, and path-independent derivative prices can be solved by, among other methods, an alternating-direction implicit finite-difference solution of the "usual" partial differential equation (PDE). Fully worked examples of solutions to these Ricatti equations and PDEs are included.

Our yield model is "affine," in the sense that there is, for each maturity τ, an affine function $Y_\tau : \mathbb{R}^n \to \mathbb{R}$ such that, at any time t, the yield of any zero-coupon bond of maturity τ is $Y_\tau(X_t)$. Indeed, ruling out singularities, essentially any n yields would serve as the factors, and given the imperfections of any model, it is an empirical issue as to which n yields will serve best as such. Likewise, because of linearity, the Markov state variables can be taken to be forward rates at given maturities, so that the model can be viewed as a multi-factor Markov parameterization of the Heath, Jarrow, and Morton (HJM) (1992) model. In fact, Frachot and Lesne (1993) have extended our model to the HJM setting.

Special cases of our model are those of Chen (1993), and Chen and Scott (1992a), Cox, Ingersoll, and Ross (CIR) (1985a) (in its multivariate form), Heston (1991), El Karoui and Rochet (1989), Jamshidian (1989, 1991, 1992), Langetieg (1980), Longstaff and Schwartz (1992), and Pennachi (1991). In all of these other earlier models, the state-variable processes are treated as "shocks" of various kinds that are not necessarily designed to be observable from the current yield curve. After solving the models for the term structure, however, the yield at any given maturity τ can be seen to be a τ-dependent affine function of the underlying state variables. Given a set of maturities equal in number with the underlying factors, one can therefore typically (that is, when the coefficient vectors defining the corresponding affine forms are linearly independent) perform a change of basis under which the state variables are yields at these various fixed maturities, as in our model. This idea has been pursued by Pearson and Sun (1994) and by Chen and Scott (1992b), who have recently and independently estimated a special case, based on a multi-factor version of Cox, Ingersoll, and Ross (1985a), by performing just such a change of variables. Our model unifies and strictly extends these affine models to the maximum possible degree, and fully exploits the idea of using yields as state variables.

Empirical studies of multi-factor models in our "affine yield" setting include those of Brown and Schaefer (1993), Chen and Scott (1992, 1993), Duffie and Singleton (1995), Frachot, Janci, and Lacoste (1992), Frachot and Lesne (1993), Heston (1991), Pearson and Sun (1994), Pennachi (1991), and Stambaugh (1988). In such parametric special cases, depending on the model specification and regularity conditions, one can generally identify the parameters of μ, σ, and R, to the extent that they affect bond prices, from cross-sectional observations of the yield curve. For example, in the one-factor CIR model, for which $r_t = X_t$ evolves according to the stochastic differential equation $dX_t = \kappa(\theta - X_t)\,dt + \gamma\sqrt{X_t}\,dW_t$, one can identify X_t, κ, θ, and γ from essentially any four distinct bond prices at time t, assuming a correct specification without measurement error. (Given the likely mis-specification of this model, this identification is not relied on in practice. Instead, it is common to use time series data and to assume fewer bond price observations at a given time than parameters, or to assume measurement error, or both.) In order to estimate the behavior of the state process X under the original probability measure P, one generally *must* resort to time-series observations, so as to capture the implied restrictions on the drift process ν. In this paper, although one of our goals is to classify a family of models that is convenient for empirical work, we are not directly concerned with estimation issues. We refer readers to the empirical studies cited above for such issues. We will restrict our attention to behavior under one particular equivalent martingale measure Q. (There may be a multiplicity of such measures in some cases, for example the case of jump-diffusions considered in Section 11.)

As with all multi-factor models, solving for all but a few types of derivative security prices is computationally intensive. We present a practical finite-difference algorithm for this purpose.

In short, we have a model specifiying simple relationships among yields, and providing term structure derivative prices, that is both computationally tractable and consistent with the absence of arbitrage. While we have not described an economy whose general equilibrium implies the behavior of the term structure appearing in our model, that is easily done along the lines of Cox, Ingersoll, and Ross (1985a,b) or Heston (1991), and adds little to what we offer.

In the model of Heath, Jarrow, and Morton (1992), as placed in a Markovian setting by Musiela (1994), the state variable is, in essence, the entire current yield curve. As such, any initial yield curve is, under regularity, consistent with the HJM model. Being in a finite-dimensional state-space setting, our model has the disadvantage that not every initial yield curve is consistent with a given paramaterization of the model. (In industry practice, this is often handled by "calibration," meaning the addition of time-dependence to the coefficients of the model in such a way as to match the given initial yield curve. That procedure has obvious disadvantages.) The disadvantage of the finite-dimensional state-space setting can also be one of its merits, for example in terms of numerical tractability. In any case, our approach of taking yields as affine factors was independently accomplished within the HJM setting by El Karoui and Lacoste (1992), taking the special Gaussian (constant volatility) case. Their work has since been extended to the stochastic volatility case by Frachot, Janci, and Lacoste (1992).

Other multi-factor term structure models include those of Litterman and Scheinkman (1988), El Karoui, Myneni, and Viswanathan (1990), Jamshidian (1993), Chan (1992), and Rogers (1995). In these models one could treat an unobserved factor as a "latent" variable that can be filtered or otherwise calibrated from observations on the yield curve.

The remainder of the paper is organized as follows. Section 2 discusses the general concepts involved in Markovian models of the yield curve. Section 3 specializes to a class of "affine factor models," in which yields are affine in some abstract state variables. It is shown that the yields are affine if, and essentially only if, the drift and diffusion functions of the stochastic differential equation for the factors are also affine. Section 4 gives conditions for existence and uniqueness of solutions to the associated stochastic differential equation. Section 5 specializes to the case in which the factors are yields at fixed maturities. Sections 6 and 7 present examples of constant and stochastic volatility versions of the yield-factor model, respectively, in which one of the factors is, for simplicity, the short rate itself. Section 8 deals with the partial differential equation (PDE) for derivative prices, providing a change of variables that orthogonalizes the diffusion so as to simplify the finite-difference solution. Sections 9 and 10 present examples of the solution to this PDE, in the stochastic and deterministic volatility cases, respectively, showing "convergence" to the theoretical solutions. Section 11 sketches an extension to the case of jump-diffusions.

2. General Factor Models of the Term Structure

For purposes of setting up the parametric model that we have in mind, we begin with the general idea of a factor model for the yield curve. Under a given complete probability space $(\Omega, \mathcal{F}, P)$ and the augmented filtration[1] $\{\mathcal{F}_t : t \geq 0\}$ generated by a standard Brownian motion W^* in $\mathbb{R}^n$, we suppose that there is a time-homogeneous Markov process X valued in some open subset D of $\mathbb{R}^n$ such that, for any times t and τ, the market value $p_{t,\tau}$ at time t of a zero-coupon bond maturing at time $t+\tau$ is given by $f(X_t, \tau)$, where $f \in C^{2,1}(D \times [0, \infty))$. The short rate process r is assumed to be defined by continuity, in that there is a measurable function $R : D \to \mathbb{R}$ defined as the limit of yields as maturity goes to zero, or

$$R(x) = \lim_{\tau \downarrow 0} \frac{-\log f(x, \tau)}{\tau}, \quad x \in D. \tag{2.1}$$

As is well understood from Harrison and Kreps (1979) and Harrison and Pliska (1981), as well as others to follow such as Ansel and Stricker (1991) to follow them, only technical regularity is required for the equivalence between the absence of arbitrage and the existence of an equivalent martingale measure, that is, a probability measure Q equivalent to P under which the price process of any security is a Q-martingale after normalization at each time t by the value $\exp\left(\int_0^t R(X_s)\, ds\right)$ of continual re-investment of interest from one unit of account held from time zero at the short rate.

Suppose that X satisfies a stochastic differential equation of the form

$$dX_t = \nu(X_t)\, dt + \sigma(X_t)\, dW_t^*, \tag{2.2}$$

where $\nu : D \to \mathbb{R}^n$ and $\sigma : D \to \mathbb{R}^{n \times n}$ are regular enough for (2.2) to have a unique (strong) solution. Additional technical regularity implies that there is a standard Brownian

[1] See, for example, Protter (1990) for definitions involving the theory of stochastic processes.

motion W in $\mathbb{R}^n$ under Q such that

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dW_t, \tag{2.3}$$

where $\mu : D \to \mathbb{R}^n$ is a function that can be calculated in terms of ν, σ, and f. General equilibrium models of this form of asset pricing behavior are given by Cox, Ingersoll, and Ross (1985b) and Huang (1987). The models in these papers are actually finite-horizon models with time-dependent coefficients, but can be extended to time-homogeneous models in an infinite-horizon setting. Our work here could be extended to time-dependent coefficients merely by notational changes and minor technical regularity. Such time-dependency, by "calibration," is standard operating procedure in trading implementations of term structure models. See, for example, Ho and Lee (1986) or Black, Derman, and Toy (1990).

Here, we are interested in choices for (f, μ, σ) that are *compatible*, in the sense that we indeed have

$$f(X_t, T-t) = E\left[\exp\left(-\int_t^T R(X_s)\,ds\right) \,\middle|\, X_t\right] \; a.s., \quad 0 \le t \le T < \infty, \tag{2.4}$$

where E denotes expectation under Q. Expression (2.4) is merely the definition of Q as an equivalent martingale measure, applied to zero-coupon bonds.

Of course, it is relatively easy to construct compatible (f, μ, σ).

For example, let μ, σ, and R be defined arbitrarily so that (2.3) and the right-hand side of (2.4) are well defined, and then let f be defined by (2.4). This is the "usual" approach in arbitrage-based term structure models, as in Dothan (1978), Vasicek (1977), Richard (1978), Black, Derman, and Toy (1990), and Hull and White (1990), among many other such models in which X is the short rate itself and R is the identity. For multivariate models, we have examples such as those of El Karoui, Myneni, and Viswanathan, (1992), Jamshidian (1993), Beaglehole and Tenney (1991), and Rogers (1995), in which X is Gauss-Markov (constant σ, affine μ) and R is a linear-quadratic form. (By "affine" μ, we mean as usual that there is a constant matrix a and a vector b such that $\mu(x) = ax + b$.) Constantinides (1992) gives a general equilibrium (representative agent) parametric model that implies this sort of linear-quadratic-gaussian behavior for short rates. There are also similar general equilibrium models, such as those of Cox, Ingersoll, and Ross (1985b), Heston (1991), Longstaff and Schwartz (1992), Nielsen and Saá-Requejo (1992), and others in which one quickly arrives at an expression such as (2.4) in which we can write $R(x) = \sum_i x_i$, where the component processes $X^{(i)}, X^{(2)}, \ldots, X^{(n)}$ are univariate processes satisfying the "CIR" equation

$$dX_t^{(i)} = (a_i + b_i X_t^{(i)})\, dt + c_i \sqrt{d_i + X_t^{(i)}}\; dW_t^i, \quad X_0^{(i)} > 0, \tag{2.5}$$

for scalar coefficients a_i, b_i, c_i, and d_i. These latter models are a special case of the model presented later in this paper.

In any case, given any candidate for the short rate process r satisfying mild regularity, it is easy to support r in a general equilibrium model based on a representative agent with, say, HARA utility and a consumption process constructed in terms of r. (See, for example, Heston (1991) and Duffie (1992, Exercise 9.3).) The available equilibrium models provide useful theoretical relationships between the term structure, preferences, technology, and macro-variables such as consumption, but have yet to add much to the practical day-to-day problems of pricing and managing the risk of fixed-income instruments. For our purposes we will follow the lead of others mentioned earlier by beginning directly with some compatible model (f, μ, σ). We are particularly interested in a class of models that is likely to be numerically and empirically tractable, and eventually models in which the state vector X_t can be treated as an observation on the term structure itself, such as intended in the first model of this sort due to Brennan and Schwartz (1979), in which the proposed factors are the short rate and the yield on a consol. (The yield on a consol is the reciprocal of its price. If one computes the price of a consol in the Brennan-Schwartz model, there is no reason to expect the result to be the reciprocal of their state variable "ℓ," which is labeled the "consol rate" by Brennan and Schwartz for expositional reasons. See Duffie, Ma, and Yong (1993) for an analysis of this issue.)

3. Affine Factor Models

We will consider a class of compatible models (f, μ, σ) with

$$f(x, \tau) = \exp[A(\tau) + B(\tau) \cdot x], \tag{3.1}$$

for which, by virtue of the maintained assumption that $f \in C^{2,1}(D \times [0, \infty))$, we know that A and B are C^1 functions on $[0, \infty)$. This parametric class of models, which we call *exponential-affine* in light of the affine relationship between yields and factors, is relatively tractable and offers some empirical advantages. In explaining the model, we will use the fact that if an affine relationship of the form $\alpha+\beta\cdot x=0$ holds for all x in some non-empty open Euclidean set, then $\alpha=0$ and $\beta=0$. We call this the "matching principle."

Since $f(x,0)=1$ for all x in D, which is an open set, (3.1) and the matching principle implies the boundary conditions

$$A(0)=0, \quad B(0)=0. \tag{3.2}$$

Since R is assumed to be well-defined by (2.1), we also know that R is an affine function on D.

Consider, for a fixed maturity date T, the zero-coupon bond price process $p_t = F(X_t,t) \equiv f(X_t, T-t)$, $t \le T$. By Ito's Lemma,

$$dp_t = \mathcal{D}F(X_t,t)\,dt + F_x(X_t,t)\sigma(X_t)\,dW_t, \tag{3.3}$$

where

$$\mathcal{D}F(x,t) = F_t(x,t) + F_x(x,t)\mu(x) + \frac{1}{2}\mathrm{tr}\left[F_{xx}(x,t)\sigma(x)\sigma(x)^\top\right].$$

We can calculate from (3.1) that

$$\begin{aligned}\mathcal{D}F(x,t) = F(x,t)\big[&-A'(T-t) - B'(T-t)\cdot x + B(T-t)\cdot\mu(x)\\ &+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} B_i(T-t)B_j(T-t)\sigma_i(x)\sigma_j(x)^\top\big].\end{aligned} \tag{3.4}$$

By (2.4), we also know that $\mathcal{D}F(X_t,t) - R(X_t)F(X_t,t)=0$. Since F is strictly positive valued, from (3.4) we have

$$\begin{aligned}0 = &-R(x) - A'(\tau) - B'(\tau)\cdot x + B(\tau)\cdot\mu(x)\\ &+\frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)\sigma_i(x)\sigma_j(x)^\top, \quad (x,\tau)\in D\times[0,\infty).\end{aligned} \tag{3.5}$$

This equation applies for all $\tau<\infty$ since T is arbitrary.

Under a mild non-degeneracy condition, (3.5) implies that μ and $\sigma\sigma^\top$ are affine functions on D. In order to see this, we can re-write (3.5) as

$$a(x,\tau) = \sum_{i=1}^{n} B_i(\tau)\mu_i(x) + \frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)\gamma_{ij}(x), \quad (x,\tau)\in D\times[0,\infty), \tag{3.6}$$

where $a(x,\tau) = R(x) + A'(\tau) + B'(\tau)\cdot x$ and where $\gamma_{ij}(x) = \sigma_i(x)\sigma_j(x)^\top$. Since R is known to be affine, for each fixed τ the function $a(\,\cdot\,,\tau)$ is affine. We let H be the function on D into $I\!R^N$, for $N = 2n + (n^2 - n)/2$, defined by

$$H(x) = (\mu_1(x), \mu_2(x), \ldots, \mu_n(x), \gamma_{11}(x), \gamma_{12}(x), \ldots, \gamma_{nn}(x))^\top\,,$$

where only the "upper-triangular" $\gamma_{ij}(x)$, those with $i \le j$, are included. We want to show that each element of H is affine in x.

We can now view (3.6) as a system of equations in τ and x of the form

$$a(x,\tau) = c(\tau)^\top H(x), \quad (x,\tau) \in D \times [0,\infty), \tag{3.7}$$

where $c : [0,\infty) \to I\!R^N$. For example, $c_1(\tau) = B_1(\tau)$ (the coeffcient of $H_1(x) = \mu_1(x)$) while $c_{n+1}(\tau) = B_1(\tau)^2/2$ (the coefficient of $\gamma_{11}(x)$.)

We can repeat (3.7) for each of any N maturities $m_1, \ldots, m_N$ to get

$$C(m_1, \ldots, m_N)H(x) = \begin{pmatrix} a(x,m_1) \\ a(x,m_2) \\ \vdots \\ a(x,m_N) \end{pmatrix}, \quad x \in D \tag{3.8}$$

where $C(m_1, \ldots, m_N)$ is the $N \times N$ matrix whose i-th row is $c(m_i)^\top$. If $C(m_1, \ldots, m_N)$ can be chosen to be non-singular, then H must be affine, as stated and proved in the following proposition, which generalizes a one-dimensional result of Brown and Schaefer (1993). Of course, for arbitrary distinct non-zero times $m_1, \ldots, m_N$, the matrix $C(m_1, \ldots, m_N)$ is non-singular except for $(B(m_1), \ldots, B(m_N))$ in a closed subset of measure zero of $I\!R^{Nn}$, which means that the affine characterization given below for $(\mu, \sigma\sigma^\top)$ is both sufficient and generically necessary for the affine yield-relationship (3.1).

Now, suppose indeed that $\mu(x)$ and $\sigma(x)\sigma(x)^\top$ are affine in x. For any fixed i, we can collect all terms in x_i from (3.5) into an expression of the form $[-B_i'(\tau) + \mathcal{B}_i(B(\tau))]x_i$, where $\mathcal{B}_i(B(\tau))$ is of the form $a + \sum_j b_j B_j(\tau) + \sum_{jk} d_{jk} B_j(\tau) B_k(\tau)$ for fixed coefficients a, b_j, and d_{jk}. That is, $\mathcal{B}_i$ is "linear-quadratic." By the matching principle, since (3.5) holds for x in an open set, we must have $-B_i'(\tau) + \mathcal{B}_i(B(\tau)) = 0$. This is true for all i and τ, giving us the differential equation

$$B'(\tau) = \mathcal{B}(B(\tau)), \quad B(0) = 0, \tag{3.8}$$

where $\mathcal{B} : I\!R^n \to I\!R^n$ is linear-quadratic. The ordinary differential (3.8) is sometimes known as a Ricatti equation.

Now, the term in (3.5) not involving x is of the form $-A'(\tau) + \mathcal{A}(B(\tau))$, where $\mathcal{A} : I\!R^n \to I\!R$ is also linear-quadratic. This term must also be identically zero in order for (3.5) to be satisfied, again by the matching principle. This implies the equation

$$A'(\tau) = \mathcal{A}(B(\tau)), \quad A(0) = 0, \tag{3.9}$$

to be solved for A, with the unique solution

$$A(\tau) = \int_0^\tau \mathcal{A}((B(s))\, ds, \tag{3.10}$$

where B solves (3.8).

There is a non-trivial issue of existence of finite solutions to Ricatti equations, since the coefficients are not Lipschitz. Solutions exist on the whole time domain for special cases such as that of Cox-Ingersoll-Ross, and, for any given particular case, they exist up to some time $\overline{T} > 0$, since (3.8) is locally Lipschitz. We implicitly assume that $\overline{T} = +\infty$ in the following proposition, but the results apply more generally by restricting $\overline{T}$.

Proposition. *Suppose (f, μ, σ) is a compatible term structure factor model and there is a finite solution to the ordinary differential equation (3.8). If μ, $\sigma\sigma^\top$, and R are affine, then f is exponential-affine. Conversely, if f is exponential-affine and there exist maturities $m_1, \ldots, m_N$ such that $C(m_1, \ldots, m_N)$ is non-singular, then μ, $\sigma\sigma^\top$, and R are affine.*

Proof: First, suppose that μ, $\sigma\sigma^\top$, and R are affine. Consider the candidate solution for f given by (3.1) for some A and B. If we can choose A and B so that (3.5) is satisfied, then the first part of the result follows. Since (3.8) has a unique solution, so does (3.9), and there is indeed a solution A and B to (3.5), implying (since f is uniquely defined) that f is exponential-affine.

Conversely, suppose that f is exponential-affine. Then R is affine. If, moreover, there exists $m_1, \ldots, m_N$ such that $C(m_1, \ldots, m_N)$, as defined above, is non-singular, then there is a unique solution $H(\cdot)$ to (3.8), which is a linear combination of affine functions, and is therefore affine. This completes the proof. ∎

It should be noted that the solution for (A, B) is not uniquely defined by the coefficients of the affine forms μ and $\sigma\sigma^\top$; it also depends on the coefficients defining R. Although there are few closed-form solutions for Ricatti equations, the solutions can be quickly computed numerically; an example is given later in the paper. For the one-dimensional cases considered by Vasicek (1977) and Cox, Ingersoll, and Ross (CIR) (1985a), there is an explicit solution for (A, B). Likewise, for the previous extensions of the CIR model in the literature, there is an explicit solution for B by virtue of the fact that f is of the form of a product $\Pi_i f_i(X^{(i)}, \tau)$, where f_i is of the form of the CIR discount bond-price function. Chen (1993) provides a 3-factor special case, distinct from the 3-factor CIR model, for which closed-form solutions are also available. For the case of μ and $\sigma\sigma^\top$ that are affine but time-dependent, the same affine yield model (3.1) applies, with (A, B) the solution of Ricatti equations with time-dependent coefficients.

4. The Affine Stochastic Differential Equation

As indicated by the last theorem, the affine class of term structure models seems to be well-behaved and offers reasonable tractability, via (3.8)-(3.9). Now we address the conditions on the domain D and the coefficients of the affine forms μ and $\sigma\sigma^\top$ under which there is indeed a unique (strong) solution to the SDE (2.3).

Without loss of generality for our purposes, we take σ to be symmetric, because for empirical issues or asset-pricing purposes there is no effect of replacing $\sigma(x)$ with a matrix square root (x by x) of $\sigma(x)\sigma(x)^\top$. The appendix shows that, if $\sigma\sigma^\top$ is affine in x, then, under non-degeneracy conditions and a possible re-ordering of indices, we can take (2.3) to be of the form

$$dX_t = (aX_t + b)\, dt + \Sigma \begin{pmatrix} \sqrt{v_1(X_t)} & 0 & \cdots & 0 \\ 0 & \sqrt{v_2(X_t)} & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \sqrt{v_n(X_t)} \end{pmatrix} dW_t, \quad X_0 \in D, \tag{4.1}$$

where $a \in I\!R^{n\times n}$, $b \in I\!R^n$, and $\Sigma \in I\!R^{n\times n}$, and

$$v_i(x) = \alpha_i + \beta_i \cdot x, \tag{4.2}$$

where, for each i, α_i is a scalar and $\beta_i \in I\!R^n$. For existence of unique solutions, coefficient restrictions apply, as indicated below.

The coefficient vectors $\beta_1, \ldots, \beta_n$ generate "stochastic volatility" unless they are all zero, in which case (4.1) defines a Gauss-Markov process. The Gauss-Markov (constant volatility) case, originally treated by Vasicek (1977) and Langetieg (1980), is reconsidered by El Karoui and Lacoste (1992) in independent work in the framework of Heath, Jarrow, and Morton (1992). This Gaussian case certainly presents no difficulty in terms of existence and uniqueness of solutions to (4.1), provided $D = I\!R^n$. With stochastic volatility, however, there is an existence issue to consider.

There are actually two delicate issues to overcome in order to assure strong solutions to (4.1). First, the diffusion function σ is not Lipschitz. Second, the volatility process $v_i(X_t)$ clearly must be non-negative for all i and t. The open domain D implied by non-negative volatilities is

$$D = \{x \in I\!R^n : v_i(x) > 0,\, i \in \{1, \ldots, n\}\}\,. \tag{4.3}$$

We must ensure that there is a unique solution to (4.1) that remains in D. For a solution X to exist, we will therefore need to assume, in effect, that for each i the volatility process $v_i(X_t)$ has a sufficiently strong positive drift on the i-th boundary segment $\partial D_i = \{x \in \overline{D} : v_i(x) = 0\}$.

Condition A. *For all i:*

(a) For all x such that $v_i(x) = 0$, $\beta_i^\top(ax + b) > \beta_i^\top \Sigma\Sigma^\top \beta_i/2$.

(b) For all j, if $(\beta_i^\top \Sigma)_j \neq 0$, then $v_i = v_j$.

Both parts of Condition A are designed to ensure strictly positive volatility, and are both effectively necessary for this purpose. Part (b) ensures that the i-th volatility term, when at zero, cannot be driven "negative" by dependence on other non-zero volatilities. (This part (b) can be relaxed to replace "$v_i = v_j$" with "$v_i = kv_j$ for some positive scalar k," but that scalar k can be absorbed into the constant matrix Σ and treated as 1 without loss of generality.) Condition A is not generically satisfied, and is a significant restriction on the model. An example satisfying Condition A (beyond the obvious Gaussian case of $\beta_i = 0$ for all i) is given later in the paper. The "stacked" univariate "square-root" processes appearing in Cox, Ingersoll, and Ross (1985a), Heston (1991), Longstaff and Schwartz (1992), and Chen and Scott (1993) all satisfy Condition A. Part (b) allows for (and goes beyond) the stacking of multivariate versions of the CIR model, each component of which is a multivarate process of the form in (4.1) with identical stochastic volatility term in each dimension.

Regarding part (a) of Condition A, Ikeda and Watanabe (1981) show that $\beta_i^\top b > \beta_i^\top \Sigma\Sigma^\top \beta_i/2$ is necessary and sufficient for zero to be an entrance boundary (that is, never hit) for a univariate process V defined by $dV_t = \beta_i \cdot b\, dt + \sqrt{V_t}\beta_i \Sigma\, dW_t$ with $V_0 > 0$. The proof, found in the appendix, of the following theorem extends this idea to the multivariate case, using part (a) of Condition A. Again, the intuition is that a sufficiently positive drift for the process $v_i(X_t)$ near the boundary where its own "volatility" is zero will ensure that this boundary is never hit. The proof is somewhat complicated by the fact that the square root function appearing in the diffusion has a derivative that approaches infinity as the stochastic volatility term $v_i(x)$ goes to zero. The reader uninterested in the details can easily skip the proof, found in the appendix, at no cost to what follows.

THEOREM. *Under Condition A, there is a unique (strong) solution X in D to the stochastic differential equation (4.1)-(4.2)-(4.3). Moreoever, for this solution X, and for all i, we have $v_i(X_t) > 0$ for all t almost surely.*

It is worth remarking that for the state process X given by this theorem, there is always a strictly positive non-constant short rate process $R(X_t)$. This follows from the Separating Hyperplane Theorem and the fact that D, as an interesection of open half-spaces, is a convex open set. For example, one could take $R(x) = \sum_i \gamma_i v_i(x)$ for non-negative γ_i.

5. The Affine Yield-Factor Models

The previous sections presented a relatively general theory of affine term structure models with abstract factors. At this point we would like to consider situations in which, for fixed maturities $\tau_1, \ldots, \tau_n$, for each i and t, we can view X_{it} as the yield at time t on a zero-coupon bond of maturity τ_i. The practical advantages of choosing factors that are yields at fixed maturities seem evident. In order for (f, μ, σ) to be an affine factor model with $f(x, \tau) = \exp(A(\tau) + B(\tau) \cdot x)$, and,

$$x_i = \frac{-1}{\tau_i} \log f(x, \tau_i), \quad x \in D, \quad i \in \{1, \ldots, n\}, \tag{5.1}$$

we need not only the initial conditions $A(0) = 0$ and $B(0) = 0$ for (3.8)-(3.9), but also, for all i:

$$A(\tau_j) = B_i(\tau_j) = 0, \quad j \neq i, \quad B_i(\tau_i) = -\tau_i. \tag{5.2}$$

We call a compatible factor model (f, μ, σ) satisfying (3.1) and (5.2) an *affine yield-factor model.*

There are two possible ways to construct an affine yield-factor model. One is to suppose from the beginning that the factors are yields, and to ensure that the coeffcients defining (f, μ, σ) are chosen so that (5.2) is satisfied. We will get to this direct approach a bit later.

The other, indirect, approach is to allow X to be the state process for an arbitrary affine factor model (f, μ, σ), and to attempt a change of variables from the original state vector X_t to a new yield state vector Y_t in $\mathbb{R}^n$ defined by

$$Y_{it} = -\frac{A(\tau_i) + B(\tau_i) \cdot X_t}{\tau_i}. \tag{5.3}$$

Provided the matrix K, whose (i, j)-element is $-B_j(\tau_i)/\tau_i$, is non-singular, we know that $X_t = K^{-1}(Y_t + k)$, where $k_i = A(\tau_i)/\tau_i$, making the change of variables possible. In this case, we can write

$$dY_t = \mu^*(Y_t)\, dt + \sigma^*(Y_t)\, dW_t, \tag{5.4}$$

where

$$\mu^*(y) = K\mu(K^{-1}(y + k)), \quad \sigma^*(y) = K\sigma(K^{-1}(y + k)), \tag{5.5}$$

which is well defined in the domain $D^* = \{Kx - k : x \in D\}$. The equivalent term structure model is (f^*, μ^*, σ^*), where

$$f^*(y, \tau) = \exp[A^*(\tau) + B^*(\tau) \cdot y], \tag{5.6}$$

for $A^*(\tau) = A(\tau) + B(\tau)^\top K^{-1}k$ and $B^*(\tau)^\top = B(\tau)^\top K^{-1}$. Clearly, (f^*, μ^*, σ^*) is an affine yield factor model.

While we have accomplished our goal indirectly, via this change of variables, for practical purposes the "covariance" function $\sigma^*(\,\cdot\,)\sigma^*(\,\cdot\,)^\top$ defined by (5.5) may be cumbersome to "calibrate" to observed volatilities or correlations, say from current or historical option-related price data, since the matrix K depends, via a solution of the Ricatti equation (3.8)-(3.9), on the original parameters defining μ and σ. There may be some practical reasons, then, to begin with an affine factor model (f, μ, σ) for which the the state vector X_t is already treated as a vector of yields at fixed maturities $\tau_1, \ldots, \tau_n$. The matrix Σ and the volatility-related coefficients α_i and β_i could be chosen directly from calibration or econometric estimation, or both. There remains, however, the question of consistency with the definition of X_{it} as the zero-coupon yield for maturity τ_i, that is, with the boundary condition (5.2). Only by adjusting the coefficients in μ or σ can we expect the solution to (3.8)-(3.9) to satisfy both the initial (zero) boundary conditions as well as the boundary conditions in (5.2). At the same time, we need to respect Condition A, of Section 4, which guarantees the existence of a solution to the stochastic differential equation defined by (μ, σ). We do not have theoretical results describing how certain coefficients can be fixed in advance and others can be adjusted to so as to achieve consistency with these various conditions. In practice, however, we have encountered no problem in fixing the coefficients in σ and then adjusting the drift coefficients so as to obtain consistency, at least in two-factor implementations. Certainly, by counting the number of equations and the number of unknowns, this success is not surprising. In the next sections we explain how to do this for 2-factor versions of the model.

6. Simple Examples

As an example to illustrate our method, we will now give more explicit treatment for special cases involving a single volatility factor or the non-stochastic volatility case: $\beta = 0$. In the latter case, the solution of the stochastic differential equation for the factors is Gaussian. The independent work of El Karoui and Lacoste (1992), is in this Gaussian setting, although they work with forward rate models in the Heath-Jarrow-Morton framework. That is, they actually take the factors to be the forward rates at certain maturities, in the sense of Heath, Jarrow, and Morton (1992). Since the yield at any maturity is affine in the factors, yield-factor and forward-rate factor models are mathematically equivalent in our setting, but HJM goes beyond this by allowing any initial term structure. El Karoui and Lacoste also provide extensive discussion of the choice of factors. (See also Frachot, Janci, and Lacoste (1992) as well as Frachot and Lesne (1993) for factor representations of the model.)

¿From this point, for simplicity, we take $X_t = (X_{0t}, X_{1t}, \ldots, X_{n-1,t})$, with $R(X_t) = X_{0t}$. That is, we take one of the factors to be the short rate itself. The slight changes in notation occasioned by this should be apparent without further comment. We also take a single stochastic volatility term; that is, $v_i(x) = v_j(x) = \alpha + \beta \cdot x$ for all i, j, and x.

While it is traditional to take one of the factors to be the short rate, there is no need for this. In fact, taking the short rate itself as a state variable can cause empirical difficulties, at least if the model is fitted to "short rate" data, which tend to have idiosyncrasies. Indeed, one may argue that the short rate itself is literally unobservable, as it is a limit of yields, rather than a yield itself.

In this special case, (3.8)-(3.9) can be written

$$\begin{cases} A'(t) = b \cdot B(t) + \frac{\alpha}{2} q(t) \\ B_0'(t) = a_0 \cdot B(t) + \frac{\beta_0}{2} q(t) - 1 \\ B_i'(t) = a_i \cdot B(t) + \frac{\beta_i}{2} q(t), \qquad i \in \{1, \ldots, n-1\}, \end{cases} \tag{6.1}$$

where

$$q(t) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \Sigma_i \Sigma_j^\top B_i(t) B_j(t).$$

For the case of determinsitic volatility, defined by $\beta = 0$, the last n equations form a simple linear system and have the standard solution:

$$B_i(\tau) = \sum_{j=0}^{n-1} \psi_{ij} \exp(\lambda_j \tau) + \psi_{in}, \quad i \in \{0, 1, \ldots, n-1\}, \tag{6.2}$$

where $\{\psi_{ij}\}$ are constants that can be easily computed and $\{\lambda_1, \ldots, \lambda_n\}$ are the n roots (assuming no multiplicity) of the characteristic equation:

$$\det \begin{pmatrix} a_{00} - \lambda & a_{01} & \cdots & a_{0,n-1} \\ a_{10} & a_{11} - \lambda & \cdots & a_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \cdots & a_{n-1,n-1} - \lambda \end{pmatrix} = 0. \tag{6.3}$$

This solution for B is then put into the first equation of (6.1) to obtain A by easy integration.

If we use only first the two factors, x_0 and x_1, we have

$$\lambda_{0,1} = \frac{1}{2} \left(a_{00} + a_{11} \pm \sqrt{\Delta} \right), \tag{6.4}$$

where $\Delta = a_{00}^2 + a_{11}^2 + 4a_{01}a_{10} - 2a_{00}a_{11}$. The constraints in (5.2) can then be explicitly written as constraints on the drift coefficients of the form:

$$
\begin{aligned}
&(a_{00} - a_{11} + \sqrt{\Delta})(a_{00} + a_{11} - \sqrt{\Delta})e^{\lambda_0 \tau_1} \\
&\qquad - (a_{00} - a_{11} - \sqrt{\Delta})(a_{00} + a_{11} + \sqrt{\Delta})e^{\lambda_1 \tau_1} - 4a_{11}\sqrt{\Delta} = 0 \qquad (6.5)\\
&a_{01}(a_{00} + a_{11} - \sqrt{\Delta})e^{\lambda_0 \tau_1} - a_{01}(a_{00} + a_{11} + \sqrt{\Delta})e^{\lambda_1 \tau_1} \\
&\qquad + 2\sqrt{\Delta}[a_{01} + \tau_1(a_{01}a_{10} - a_{00}a_{11})] = 0. \qquad (6.6)
\end{aligned}
$$

This deterministic volatility example is extended in Section 10, where we offer explicit and numerical solutions for bond option prices by adapting to our model the results of Jamshidian (1989, 1991), El Karoui and Rochet (1989), and others.

7. Two-Factor Stochastic-Volatility Model

We now concentrate on the two-factor case.

First, we consider the coefficient restriction required for non-negativity of volatility. In this case, the "hyperplane" defining zero volatility is given by

$$H = \{(x_0, x_1) : \alpha + \beta_0 x_0 + \beta_1 x_1 = 0\}. \qquad (7.1)$$

Without loss of generality if $\beta_1 \neq 0$, we take $\beta_1 = 1$, so that on H we have $x_1 = -(\alpha + \beta_0 x_0)$. On H, the drift function for $V_t = \alpha + \beta_0 X_{0t} + \beta_1 X_{1t}$ is therefore

$$
\begin{aligned}
\beta_0\mu_0(x) + \beta_1\mu_1(x) &= \beta_0[b_0 + a_{00}x_0 - a_{01}(\alpha + \beta_0 x_0)] + b_1 + a_{10}x_0 - a_{11}(\alpha + \beta_0 x_0) \\
&= k_0 + k_1 x_0,
\end{aligned}
$$

where

$$
\begin{aligned}
k_0 &= \beta_0 b_0 - \alpha\beta_0 a_{01} + b_1 - \alpha a_{11} \\
k_1 &= \beta_0 a_{00} - \beta_0^2 a_{01} + a_{10} - \beta_0 a_{11}.
\end{aligned}
$$

In this case, an affine yield-factor model calls for b, a, and Σ in a manifold satisfying (6.1), (5.2), and

$$k_0 > 0 \quad \text{and} \quad k_1 = 0. \qquad (7.2)$$

We give an example in Section 9.

8. Finite-Difference Solution of Derivative Asset Prices

By the definition of an equivalent martingale measure, an asset defined by a payoff u at time T has a price at any time $t < T$ given by

$$E\left[\exp\left(-\int_t^T R(X_s)\,ds\right)u \;\Big|\; \mathcal{F}_t\right].$$

If u is a random variable that is a measurable function of the term structure at time T, then (since the term structure is itself a measurable function of the state variables X_T), we can write $u = g(X_T)$, and express the price in the form

$$F(X_t, t) = E\left[\exp\left(-\int_t^T R(X_s)\,ds\right)g(X_T) \;\Big|\; X_t\right]. \tag{8.1}$$

Under mild regularity conditions (see, for example, Friedman (1975)), the unique solution to (8.1) satisfies the PDE

$$\mathcal{D}F(x,t) - R(x)F(x,t) = 0, \quad x \in D, \tag{8.2}$$

where $\mathcal{D}F(x,t)$ is as defined by (3.3), with the boundary condition

$$F(x,T) = g(x), \quad x \in D. \tag{8.3}$$

There are well known finite-difference algorithms for solving a parabolic PDE of this form. In order to simplify the numerical solution in the two-factor case described in the previous section, it is convenient to make the change of variables:

$$\begin{aligned} y &= \frac{1}{1 + k\sqrt{\alpha + \beta_0 x_0 + \beta_1 x_1}} \\ z &= \arctan[h(\theta_1 x_1 - \theta_0 x_0)], \end{aligned} \tag{8.4}$$

where $\theta_0 = \beta_1\Sigma_1\Sigma_1^\top + \beta_0\Sigma_0\Sigma_1^\top$ and $\theta_1 = \beta_0\Sigma_0\Sigma_0^\top + \beta_1\Sigma_0\Sigma_1^\top$. It is easy to see that $0 \le y \le 1$ and $-\pi/2 \le z \le \pi/2$.

The inverse of this transformation is given by

$$\begin{aligned} x_0 &= \xi_0 + \eta_0\frac{(1-y)^2}{k^2y^2} + \zeta_0\frac{\tan(z)}{h} \\ x_1 &= \xi_1 + \eta_1\frac{(1-y)^2}{k^2y^2} + \zeta_1\frac{\tan(z)}{h}, \end{aligned} \tag{8.5}$$

where

$$\xi_0 = -\frac{\alpha\theta_1}{\beta_0\theta_1 + \beta_1\theta_0}$$

$$\xi_1 = -\frac{\alpha\theta_0}{\beta_0\theta_1 + \beta_1\theta_0}$$

$$\eta_0 = \frac{\theta_1}{\beta_0\theta_1 + \beta_1\theta_0}$$

$$\eta_1 = \frac{\theta_0}{\beta_0\theta_1 + \beta_1\theta_0}$$

$$\zeta_0 = -\frac{\beta_1}{\beta_0\theta_1 + \beta_1\theta_0}$$

$$\zeta_1 = \frac{\beta_0}{\beta_0\theta_1 + \beta_1\theta_0}.$$

Then, (8.2) is written, for $\tilde{F}(y,z,t) \equiv F(x_0,x_1,t)$, in the form

$$-\mu\tilde{F} + \mu_t\tilde{F}_t + \mu_y\tilde{F}_y + \mu_z\tilde{F}_z + \sigma_y\tilde{F}_{yy} + \sigma_z\tilde{F}_{zz} = 0, \tag{8.6}$$

where

$$\begin{aligned}
\mu_t =& k^2y^2(1-y)\cos(z)\\
\mu \ =& k^2\xi_0y^2(1-y)\cos(z) + \eta_0(1-y)^3\cos(z) + k^2\zeta_0y^2(1-y)\frac{\sin(z)}{h}\\
\mu_y =& -\frac{k^2y^3}{2}\left(\mu_{y0}k^2y^2\cos(z) + \mu_{y1}(1-y)^2\cos(z) + \mu_{y2}k^2y^2\frac{\sin(z)}{h}\right)\\
&+\frac{k^4(\Sigma_0\Sigma_0^\top\beta_0^2 + \Sigma_1\Sigma_1^\top\beta_1^2 + 2\Sigma_0\Sigma_1^\top\beta_0\beta_1)}{8}y^5(3-2y)\cos(z)\\
\mu_z =& -h(1-y)\cos^2(z)\left(\mu_{z0}k^2y^2\cos(z) + \mu_{z1}(1-y)^2\cos(z) + \mu_{z2}k^2y^2\frac{\sin(z)}{h}\right)\\
&-h^2(\Sigma_0\Sigma_0^\top\theta_0^2 + \Sigma_1\Sigma_1^\top\theta_1^2 - 2\Sigma_0\Sigma_1^\top\theta_0\theta_1)(1-y)^3\cos^4(z)\sin(z)\\
\sigma_y =& \frac{k^4(\beta_0^2\Sigma_0\Sigma_0^\top + 2\beta_0\beta_1\Sigma_0\Sigma_1^\top + \beta_1^2\Sigma_1\Sigma_1^\top)}{8}y^6(1-y)\cos(z)\\
\sigma_z^2 =& \frac{h^2(\theta_0^2\Sigma_0\Sigma_0^\top - 2\theta_0\theta_1\Sigma_0\Sigma_1^\top + \theta_1^2\Sigma_1\Sigma_1^\top)}{2}(1-y)^3\cos^5(z),
\end{aligned}$$

with

$$\begin{aligned}
\mu_{y0} =& \beta_0 b_0 + \beta_1 b_1 + \beta_0 a_{00}\xi_0 + \beta_1 a_{10}\xi_0 + \beta_0 a_{01}\xi_1 + \beta_1 a_{11}\xi_1 \\
\mu_{y1} =& \beta_0 a_{00}\eta_0 + \beta_1 a_{10}\eta_0 + \beta_0 a_{01}\eta_1 + \beta_1 a_{11}\eta_1 \\
\mu_{y2} =& \beta_0 a_{00}\zeta_0 + \beta_1 a_{10}\zeta_0 + \beta_0 a_{01}\zeta_1 + \beta_1 a_{11}\zeta_1 \\
\mu_{z0} =& \theta_0 b_0 - \theta_1 b_1 + \theta_0 a_{00}\xi_0 - \theta_1 a_{10}\xi_0 + \theta_0 a_{01}\xi_1 - \theta_1 a_{11}\xi_1 \\
\mu_{z1} =& \theta_0 a_{00}\eta_0 - \theta_1 a_{10}\eta_0 + \theta_0 a_{01}\eta_1 - \theta_1 a_{11}\eta_1 \\
\mu_{z2} =& \theta_0 a_{00}\zeta_0 - \theta_1 a_{10}\zeta_0 + \theta_0 a_{01}\zeta_1 - \theta_1 a_{11}\zeta_1.
\end{aligned}$$

The reduction of (8.2) to (8.6) has two main advantages. First, we have converted the coordinates x_0 and x_1, which in general take any real values, to coordinates y and z that take values in compact sets. An evenly spaced grid over these compact sets implies a concentration of grid points in the original variables that can be controlled for accuracy of the solution, placing greater grid density near more frequently encountered rates. Second, we have orthogonalized the system so that an alternating-directions finite-difference method can be applied, given the absence in (8.6) of cross partial derivatives in the two new space variables, y and z. For the alternating directions implicit method, see Ames (1977) and Press *et al* (1988).

9. Stochastic Volatility Example

In this section, we give an example of the two-factor model discussed in last section.

We first solve the two-dimensional ODE

$$\begin{cases} B_0'(t) = a_{00}B_0(t) + a_{10}B_1(t) + \dfrac{\beta_0}{2}q(t) - 1 \\ B_1'(t) = a_{01}B_0(t) + a_{11}B_1(t) + \dfrac{\beta_1}{2}q(t), \end{cases} \tag{9.1}$$

where

$$q(t) = \Sigma_0\Sigma_0^\top B_0^2(t) + \Sigma_1\Sigma_1^\top B_1^2(t) + 2\Sigma_0\Sigma_1^\top B_0(t)B_1(t), \tag{9.2}$$

subject to the initial conditions $B_0(0) = B_1(0) = 0$, using the fourth-order Runge-Kutta method. The resulting solution (B_0, B_1) depends on the parameter vector (a, Σ). Then, fixing Σ, we use a Newton-Raphson algorithm to solve for a so as to match the consistency conditions $B_0(\tau_1) = 0$ and $B_1(\tau_1) = -\tau_1$ of (5.2). As is well known, the success of the Newton-Raphson method in multi-dimensions depends critically on the accuracy of the first guess. We suggest that one finds the solution to the deterministic volatility case, and use that as the first guess here. For the fourth-order Runge-Kutta ODE solution method and Newton-Raphson search method, please see Press *et al* (1988).

Given the solution of (9.1) for B, we numerically integrate

$$A'(t) = b_0 B_0(t) + b_1 B_1(t) + \frac{\alpha q(t)}{2}, \quad A(0) = 0, \tag{9.3}$$

to obtain A, and then choose b_0 and b_1 so as to match the consistency condition $A(\tau_1) = 0$.

As an example, we fix the parameters: $b_0 = -0.1$, $a_{01} = 0.76$, $\Sigma_0 \Sigma_0^\top = \Sigma_1 \Sigma_1^\top = 0.9$, $\Sigma_1 \Sigma_2^\top = 0.7$, $\alpha = 0.05$, $\beta_0 = -1.08$ and $\beta_1 = 1$. We use the Newton-Raphson method to find a_{00} and a_{11} satisfying the consistency conditions. The rule used to terminate the Newton-Raphson iteration is to stop when both $\mid B_0(\tau_1) \mid < 0.0001$ and $\mid B_1(\tau_1) + \tau_1 \mid < 0.0001$. Table 1 presents an example of the results. The grid size is the reciprocal of the step size used to numerically solve the ODE. The coefficient a_{10} is calculated by setting $k_1 = 0$ in (7.2), while b_1 is calculated by setting $A(\tau_1) = 0$. The resulting set of parameters satisfies Condition A for the existence and uniqueness of solutions to the SDE (4.1)-(4.2)-(4.3).

Table 1. Example Parameter Solution

grid size n	a_{00}	a_{11}	a_{10}	b_1
10001	-0.6419	3.7377	-3.900	0.1278
20001	-0.6418	3.7374	-3.900	0.1278

Using the parameters obtained above, we use the alternating direction implicit (ADI) method to solve the PDE (8.2) in the form of (8.6). In relative terms, solving for the parameters b and a is a much faster procedure than solving this PDE for given parameters. As is well-known, there is no general theory guaranteeing the convergence of the ADI algorithm when applied to a specific problem. In our case, this method is in fact divergent near some boundaries, perhaps due to the rapid change of the value of $\tilde{F}$ near those boundaries. (Indeed, $\tilde{F}$ is infinite along some boundaries because x_0 and x_1 can be negative in this parameterization of our model.) In order to restore convergence, we apply the ADI algorithm on the domain: $y \in [\delta_1, 1-\delta_2]$ and $z \in [-\pi/2+\delta_3, \pi/2-\delta_4]$, where δ_1, δ_2, δ_3 and δ_4 are small non-negative numbers. By appropriately choosing these small numbers, we indeed obtain convergence as shown below. The computation time can in principle be improved by "hopscotch" methods, which alternate implicit and explicit steps.

In order to examine the precision of this method, Table 2 shows numerical results for the price of a zero-coupon bond with 1 unit of time to maturity. The exact result is given by $\exp(-x_1)$. In the parenthesis, we also give the numerical solution to the PDE for the price for an American call on this bond maturing at time 0.5, with strike price 0.9. We have chosen $h = k = 12.5$, $\delta_1 = 0.01$, $\delta_2 = \delta_3 = 0$, and $\delta_4 = 0.01$.

Table 2. Example Finite-Difference PDE Solution of Bond and Bond Option Prices
(Bond Option Prices in Parentheses)

		Grid Size N			Exact
short rate (x_0)	long rate (x_1)	111	221	331	∞
0.1070	0.1584	0.8464 (0.0578)	0.8531 (0.0610)	0.8532 (0.0606)	0.8535
0.0336	0.0791	0.9179 (0.0888)	0.9251 (0.0926)	0.9246 (0.0922)	0.9240
0.0710	0.0593	0.9411 (0.0879)	0.9420 (0.0886)	0.9421 (0.0883)	0.9424

10. Deterministic Volatility Bond Option Pricing Example

In the case of a short rate process that can be viewed as a component of a multivariate Gauss-Markov process, Jamshidian (1989, 1991), El Karoui and Rochet (1989), and others have computed the prices of bond options explicitly. In this setting, we can use our results to restrict the coefficients of the Gauss-Markov process so that the state variables can be taken to be yields. We thereby obtain a convenient example in which bond option prices can be computed in terms of the yields at the basis maturities, and can thus verify the accuracy of our numerical solution for option prices against the explicit solution.

For our example, we take the two-factor deterministic volatility yield-factor model, with $\alpha = 1$ and $\beta_0 = \beta_1 = 0$. Under the variable transformation

$$\begin{cases} y = \arctan(kx_0) \\ z = \arctan[h(\Sigma_0\Sigma_0^\top x_1 - \Sigma_0\tilde{\sigma}_1^\top x_0)], \end{cases} \tag{10.1}$$

the PDE (8.2) can be written in the form of (8.6), with

$$\begin{aligned}
\mu_t =& kh\cos(y)\cos(z) \\
\mu =& h\sin(y)\cos(z) \\
\mu_y =& kh\cos^2(y)\cos(z)[kb_0\cos(y) + a_{00}\sin(y) - k^2\Sigma_0\Sigma_0^\top\cos^2(y)\sin(y)] \\
&\quad + \frac{ka_{01}}{\Sigma_0\Sigma_0^\top}\cos^2(y)[h\Sigma_0\Sigma_1^\top\sin(y)\cos(z) + k\cos(y)\sin(z)] \\
\mu_z =& kh^2\cos(y)\cos^3(z)[b_1\Sigma_0\Sigma_0^\top - b_0\Sigma_0\Sigma_1^\top + h\cos(z)\sin(z)\Sigma_0\Sigma_0^\top((\Sigma_0\Sigma_1^\top)^2 - (\Sigma_0\Sigma_0^\top)(\Sigma_1\Sigma_1^\top))] \\
& + h^2\sin(y)\cos^3(z)(a_{10}\Sigma_0\Sigma_0^\top - a_{00}\Sigma_0\Sigma_1^\top) \\
& + \frac{h}{\tilde{\sigma}_0\Sigma_0^\top}(a_{11}\Sigma_0\Sigma_0^\top - a_{01}\Sigma_0\Sigma_1^\top)(h\Sigma_0\Sigma_1^\top\sin(y)\cos(z) + k\cos(y)\sin(z)) \\
\sigma_y =& \frac{k^3h\Sigma_0\Sigma_0^\top}{2}\cos^5(y)\cos(z) \\
\sigma_z =& \frac{kh^3\Sigma_0\Sigma_0^\top[(\Sigma_0\Sigma_0^\top)(\Sigma_1\Sigma_1^\top) - (\Sigma_0\Sigma_1^\top)^2]}{2}\cos(y)\cos^5(z).
\end{aligned}$$

For the sake of convergence, we restrict ourself to the domain $y \in [-\pi/2+\delta_1, \pi/2-\delta_2]$ and $z \in [-\pi/2+\delta_3, \pi/2-\delta_4]$.

By applying the results of Jamshidian (1989, 1991), it can be shown that the price at time t of a European call option on a zero-coupon bond that pays \$1 at time T, with strike price K and expiration time $\tau^* < T$, is given by:

$$C(X_t,t) = f(X_t, T-t)N\left(\Lambda(X_t,t) + \frac{\sigma^*(t)}{2}\right) - Kf(X_t,\tau^*-t)N\left(\Lambda(X_t,t) - \frac{\sigma^*(t)}{2}\right), \tag{10.2}$$

where

$$\Lambda(X_t,t) = \frac{1}{\sigma^*(t)}\log\left(\frac{f(X_t,T-t)}{Kf(X_t,\tau^*-t)}\right)$$

$N(\cdot)$ denotes the cumulative standard normal distribution, and σ^* is the function on $[0, \tau^*]$ given by $\sigma^*(t)^2 = \int_t^{\tau^*} H(s)\, ds$, where

$$H(s) = \Sigma_0 \Sigma_0^\top [B_0(T-s) - B_0(\tau^* - s)]^2 + \Sigma_1 \Sigma_1^\top [B_1(T-s) - B_1(\tau^* - s)]^2 + 2\Sigma_0 \Sigma_1^\top [B_0(T-s) - B_0(\tau^* - s)][B_1(T-s) - B_1(\tau^* - s)],$$

with $B_0(\cdot)$ and $B_1(\cdot)$ as given in (6.2). The option pricing formula (10.2) is a version of the Black-Scholes (1973) formula.

For our numerical example, we take:

$b_0 = -0.1$, $b_1 = 0.6453$, $a_{00} = -0.9651$, $a_{01} = 0.59$, $a_{10} = -3.21$, $a_{11} = 3.5802$, $\Sigma_0 \Sigma_0^\top = \Sigma_1 \Sigma_1^\top = 1.0$, $\Sigma_0 \Sigma_1^\top = 0.6$, $k = h = 12.5$, $\delta_1 = \delta_3 = 0.01$, and $\delta_2 = \delta_4 = 0$.

Table 3 shows the computed prices of a European Call with strike price $K = 0.9$, option expiration time $\tau^* = 0.5$, and bond maturity $T = 1.0$.

Table 3. Bond and Bond Option Prices with Deterministic Volatlity

(Bond Option Prices in Parentheses)

		Grid Size n			Exact
short rate (x_0)	long rate (x_1)	111	221	331	∞
0.1106	0.1751	0.8490 (0.1158)	0.8376 (0.1060)	0.8403 (0.1065)	0.8394 (0.1065)
0.05849	0.03509	1.0001 (0.1761)	0.9671 (0.1677)	0.9668 (0.1681)	0.9655 (0.1679)
0.02635	0.07344	0.9493 (0.1520)	0.9281 (0.1415)	0.9309 (0.1424)	0.9292 (0.1421)

11. Jump-Duffusion State Process

Because of the possibility of sudden changes in perceptions of future interest rates, one may wish to allow for "surprise" jumps in the state vector X. For example, one can maintain the affine yield-factor model with a standard jump-diffusion model for X based on the infinitesimal generator $\mathcal{D}^*$ defined by

$$\mathcal{D}^* F(x,t) = \mathcal{D}F(x,t) + \lambda(x) \int_D [F(x+z,t) - F(x,t)]\, d\nu(z), \tag{11.1}$$

where $\mathcal{D}$ is the diffusion generator defined by (3.3), $\lambda : D \to I\!R_+$ is an affine function determining the arrival intensity $\lambda(X_t)$ of jumps in X at time t, and where ν is a fixed probability measure on $I\!R^n$ defining the distribution of jumps. As before the zero-coupon bond price with maturity T has a price $F(X_t, t)$ at time t, where, under technical regularity, F solves the PDE

$$\mathcal{D}^* F(x,t) - R(x)F(x,t) = 0 \tag{11.2}$$

with the boundary condition

$$F(x,T) = 1. \tag{11.3}$$

With μ, $\sigma\sigma^\top$, ρ, and λ all affine functions on the state space D, the PDE (11.2)-(11.3) is, under regularity, solved by usual exponential affine form

$$F(x,t) = \exp\left[a(T-t) + b(T-t)\cdot x\right], \tag{11.4}$$

where $a : [0,T] \to I\!R$ and $b : [0,T] \to I\!R^n$ are solutions of ordinary differential equations that are easily computed numerically in many cases. It is convenient, as one can see from substituting (11.4) into (11.2) in order to derive the ODE for b, to choose a distribution ν whose Laplace transform $\theta(\,\cdot\,)$ is known explicitly, so as to avoid a numerical computation of the term $\theta(b(T-t))$. Combinations of exponential, binomial, degenerate (fixed jump size), and gaussian distributions are convenient, although one must take care to choose a distribution that ensures that the state process X, jumping from any point in the state space D, remains in the state space. For a Gaussian special case (in which $\sigma\sigma^\top$ is constant and ν is a Gaussian distribution on $D = I\!R^n$), closed-form solutions are given by Das (1993). Closed form solutions are also available when one chooses $\mathcal{D}$ to the the generator associated with the multi-factor CIR model, and takes ν to be a product of n exponential distributions.

By changing the boundary condition (11.3) to one appropriate for a given derivative payoff, one can also value the derivative security. Numerical solution of the PDE by finite difference is relatively straightforward, although the usual staircase algorithm for inverting the implicit difference step is not directly applicable with non-degenerate jump distributions. We have successfully implemented a numerical algorithm for option valuation with exponential jump distributions in 2-dimensional special cases.

Also, with jumps, it may be impossible to perfectly hedge a given claim with fewer positions in other claims than the cardinality of the support of the jump distribution ν.

Appendix: SDE Results

This appendix addresses the form and existence of "affine stochastic differential equations," those in some state space $D \subset I\!R^n$, of the form

$$dX_t = \mu(X_t)\, dt + \sigma(X_t)\, dW_t, \quad X_0 \in D, \tag{A.1}$$

where $\mu : D \to I\!R^n$ and $\sigma\sigma^\top : D \to \mathcal{M}_s$ are affine, taking $\mathcal{M}$ to denote the space of real $n \times n$ matrices, and $\mathcal{M}_s \subset \mathcal{M}$ to denote the subset of symmetric matrices.

Since $\Theta = \sigma\sigma^\top$ is affine, for any i and j we have $\Theta_{ij}(x) = a_{ij} + b_{ij} \cdot x$, for some a_{ij} in $I\!R$ and b_{ij} in $I\!R^n$. For each i in $L = \{i : b_{ii} \neq 0\}$, the affine space $A_i \subset I\!R^n$ of roots to the equation $a_{ii} + b_{ii} \cdot x = 0$ is an $(n-1)$-dimensional manifold defining points that, if in the state space, would be associated with zero "instantaneous" variance of the changes in the state process X.

We will fix a particular "canonical" state space $S \subset I\!R^n$. Since the diagonal elements of $\Theta(x)$ must be non-negative for all x, and are affine in x, we know that S is contained by the intersection of half-spaces $\widehat{S} = \{x : \Theta_{ii}(x) \geq 0, \quad i \in L\}$. In fact, up to closure, it is reasonable to suppose that $S = \widehat{S}$, since a point x in the boundary of S that is *not* in the boundary of $\widehat{S}$ is in the interior of $\widehat{S}$. At such a starting point, (barring degeneracies) the state process X would exit from S. We therefore take the canonical state space S to be $\widehat{S}$.

Allowing for the possibility that $A_i = A_j$ for some $i \neq j$, we can always choose some minimal subset $K \subset L$ such that $\widehat{S} = \{x : \Theta_{ii}(x) \geq 0, i \in K\}$.

NON-DEGENERACY OF Θ. *The set $\{b_{ii} : i \in K\} \subset I\!R^n$ is linearly independent.*

For example, non-degeneracy rules out parallel boundaries for the state space S, which is ruled out in any case by consideration of existence of solutions to the SDE for X, unless two of the co-ordinate processes X_i and X_j are scalings of each other. Under non-degeneracy, the sub-manifold $\widehat{A}_i = A_i \bigcap S$ is also $(n-1)$-dimensional. The boundary of S is $\bigcup_{i \in K} \widehat{A}_i$.

A *strip* is a set of the form $\{x \in I\!\!R^n : \; c \leq u(x) \leq d\} \subset I\!\!R^n$, for some $c \in I\!\!R$, $d \in I\!\!R$, and linear $u : I\!\!R^n \to I\!\!R$.

LEMMA A.1. *If Θ is non-degenerate then S is not contained by a strip.*

PROOF: Suppose not. Then there exists some linear $u : I\!\!R^n \to I\!\!R$ such that $c \leq u(x) \leq d$ for all x in S. Let y be a non-singular linear transformation of x with $y_1 = u(x)$. (That is, we pick some invertible linear $Y : I\!\!R^n \to I\!\!R^n$ such that $Y_1(x) = u(x)$, $x \in I\!\!R^n$. Throughout, we write "y" for "$Y(x)$," for any typical point x in $I\!\!R^n$.) For each $i \in K$, we have $b_{ii} \cdot x = \tilde{b}_i \cdot y$ for some $\tilde{b}_i$ in $I\!\!R^n$. It follows from non-degeneracy that $\{\tilde{b}_i : i \in K\}$ is linearly independent. There exists some $i \in K$ and some $\hat{y} \in I\!\!R^n$ with $\hat{y}_1 \neq 0$ such that $\tilde{b}_j \cdot \hat{y} = 0$ for all $j \neq i$ and $\tilde{b}_i \cdot \hat{y}_i \geq 0$. Hence, for any $y \in Y(S)$, $y + \hat{y} \in Y(S)$. Then, in order to have $y_1 \geq c$, $\hat{y}_1$ must be positive. But in order to have $y_1 \leq d$, $\hat{y}_1$ must be negative, a contradiction. ▮

We will say that a result applies to Θ "up to a re-ordering of indices" if the result applies after replacing Θ with a the function $x \mapsto \left(\Theta(x)_{\pi(i),\pi(j)}\right)$, for some permutation $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$.

Since we can, without loss of generality for our purposes, replace σ with any measurable "square root" of Θ, it is without loss of generality that we suppose $\sigma(x)$ to be symmetric for all x.

LEMMA A2. *If Θ is non-degenerate, then, up to a re-ordering of indices,*

$$\Theta(x) = \begin{pmatrix} B_1 u_1(x) & 0 & \cdots & 0 \\ 0 & B_2 u_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_M u_M(x) \end{pmatrix}, \quad x \in S, \qquad (A.2)$$

where $1 \leq M \leq n$, and for $i \in \{1, 2, \ldots, M\}$, B_i is an $N_i \times N_i$ positive semi-definite symmetric matrix, with $\sum_i N_i = n$, and where $u_1, \ldots, u_M$ are affine on $I\!\!R^n$ into $I\!\!R$, with linear components that are pairwise linearly independent.

PROOF: Because $\sigma(x)$ is symmetric, we have

$$\Theta_{ij}(x) = \sum_{k=1}^{n} \sigma_{ik}(x)\sigma_{kj}(x) = \sum_{k=1}^{n} \sigma_{ik}(x)\sigma_{jk}(x).$$

Especially,

$$\Theta_{ii}(x) = \sum_{k=1}^{n} \sigma_{ik}^2(x).$$

Hence, for x such that $\Theta_{ii}(x) = 0$, we must have $\sigma_{ik}^2(x) = 0$ for all k and thus $\Theta_{ij}(x) = 0$ for all j. ¿From this, we will show that $\Theta_{ij}(\,\cdot\,)$ is proportional to $\Theta_{ii}(\,\cdot\,)$, proving the result. There are two possible cases.

1. Suppose $\Theta_{ii}(\,\cdot\,)$ is not a constant. By the above reasoning and non-degeneracy, both Θ_{ii} and Θ_{ij} are zero everywhere on $\widehat{A}_i$, which is a relatively open subset of an $(n-1)$-dimensional affine space, A_i. We can treat A_i as a translation by some (possibly zero) scaling of b_{ii} of the linear subspace M_i orthogonal to b_{ii}. Since $\widehat{A}_i$ is relatively open and Θ_{ij} is zero everywhere on $\widehat{A}_i$, b_{ij} must also be orthogonal to M_i, and thus $b_{ij} = k_{ij} b_{ii}$ for some constant k_{ij}. We now have $a_{ii} + b_{ii} \cdot x = 0 = a_{ij} + k_{ij} b_{ii} \cdot x$ for all x in $\widehat{A}_i$. This can only be true if $a_{ij} = k_{ij} a_{ii}$. Thus, for some constant scalar k_{ij} (possibly zero), we have $\Theta_{ij} = k_{ij}\Theta_{ii}$.
2. Suppose $\Theta_{ii}(\,\cdot\,)$ is constant. In this case, $\Theta_{ij}(x)$ must also be a constant. If not, the sub-matrix
$$\begin{pmatrix} \Theta_{ii}(x) & \Theta_{ij}(x) \\ \Theta_{ji}(x) & \Theta_{jj}(x) \end{pmatrix}$$
cannot be semi-positive definite, shown as follows. There are two sub-cases to consider. If Θ_{jj} is constant, and Θ_{ij} is not, the fact that S is not contained by a strip implies failure of positive-semi-definiteness. If Θ_{ij} is not constant, then, from case (1), $\Theta_{ij}(x) = \Theta_{ji}(x)$ can be written as $k_{ji}\Theta_{jj}(x)$ for some constant scalar k_{ji}. The determinant of the above sub-matrix is then $\Theta_{ii}(x)\Theta_{jj}(x) - k_{ji}^2\Theta_{jj}^2(x)$. This determinant can be arbitrarily negative because S is not contained by a strip. This, however, contradicts the positive-semi-definite nature of $\Theta(x)$. ▮

PROPOSITION A.1. *Suppose Θ is non-degenerate, and there is some $\tilde{x} \in S$ such that $\Theta(\tilde{x})$ is positive-definite. Then there exists a non-singular constant matrix Q such that*

$$Q\Theta(x)Q^\top = \begin{pmatrix} v_1(x) & 0 & \cdots & 0 \\ 0 & v_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_n(x) \end{pmatrix}, \quad x \in I\!R^n,$$

where, for each i, $\tilde{v}_i : I\!R^n \to I\!R$ is affine.

PROOF: We can always write $\Theta(x) = A + \Lambda(x)$ where $A \in \mathcal{M}_s$ and Λ is of the form $(A.2)$, for linear u_i. There exists some non-singular constant matrix P such that $PAP^\top$ is diagonal. Since $P\Lambda(x)P^\top$ is symmetric and linear in x, it must have the form given by the right-hand side of (A.2), for linear u_i. Therefore, up to a re-ordering of indices, we have the block-diagonal form:

$$P\Theta(x)P^\top = \begin{pmatrix} A_1 + B_1 u_1(x) & 0 & \cdots & 0 \\ 0 & A_2 + B_2 u_2(x) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_M + B_M u_M(x) \end{pmatrix}, \quad x \in I\!R^n$$

where, for each block i, A_i is diagonal and B_i is symmetric. Consider a particular diagonal block i. For some $\tilde{x}$, as assumed, $A_i + B_i u_i(\tilde{x})$ is positive-definite. By a result found in Hohn (1964), there exists some non-singular matrix Q_i of the dimensions of A_i and B_i such that $Q_i(A_i + B_i u_i(\tilde{x}))Q_i^\top$ is the identity matrix and $Q_i B_i Q_i^\top$ is diagonal. Noticing that

$$A_i + B_i u_i(x) = A_i + B_i u_i(\tilde{x}) + B_i(u_i(x) - u_i(\tilde{x})),$$

we can let

$$Q = \begin{pmatrix} Q_1 & 0 & \cdots & 0 \\ 0 & Q_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_K \end{pmatrix} P.$$

Since a diagonal matrix is diagonal even after a re-ordering of indices, we have the result. ▮

COROLLARY A.1. *Under the assumptions of the Proposition,*

$$\sigma(x) = \Sigma \begin{pmatrix} \sqrt{v_1(x)} & 0 & \cdots & 0 \\ 0 & \sqrt{v_2(x)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{v_n(x)} \end{pmatrix}, \quad x \in S, \qquad (A.4)$$

where Σ is a non-singular matrix and $v_1, \ldots, v_n$ are affine functions.

PROOF: From Proposition A.1, there exists a non-singular matrix Q such that $Q\Theta(x)Q^\top = \Psi(x)$ for all x, where $\Psi(x)$ is diagonal for all x, and affine in x. Let $v_i(x) = \Psi_{ii}(x)$ and $\Sigma = Q^{-1}$. The conclusion follows immediately. ▮

This implies another characterization, as follows.

COROLLARY A.2. *Θ has the properties assumed in Proposition A.1 if and only if*

$$\Theta(x) = \sum_{i=1}^{n} V_i V_i^{\top} \tilde{w}_i(x), \tag{A.5}$$

where the vectors $V_1, \ldots, V_n$ are linearly independent in $\mathbb{R}^n$, the functions $\tilde{w}_1, \ldots, \tilde{w}_n$ are affine on $\mathbb{R}^n$ and non-negative on S, and the set $\{x : \tilde{w}_i \geq 0, \quad i \in \{1, \ldots, n\}\}$ has an interior point.

We have by now characterized Θ and $\sigma(\cdot)$ under the implicit assumption that the state space D is of the form taken for S, that is, a closed intersection of half-spaces. In fact, we can and do take D to be the interior of S, and apply conditions (Condition A of Section 4) that prevent the boundary of S from being hit. This is the focus of the remainder of this appendix.

In order to state the comparison Lemma used in the proof of the Theorem in Section 4, we record the following property of a diffusion function.

YAMADA CONDITION. *A function $\sigma : \mathbb{R}_+ \to \mathbb{R}$ satisfies the Yamada Condition if bounded and measurable, and if there exists a function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$, strictly increasing, continuous, with $\rho(0) = 0$, $\int_{0+}^{1} \rho(u)^{-2}\, du = +\infty$, and $|\sigma(u) - \sigma(v)| \leq \rho(|u - v|)$ for all u and v.*

For example, σ satisfies the Yamada condition if $\sigma(u) = \min(\sqrt{u}, k)$, for some constant k.

LEMMA A.3. *Suppose that Z is a standard Brownian motion, σ satisfies the Yamada condition, and $\mu : \mathbb{R} \to \mathbb{R}$ is Lipschitz. Then there is a unique (strong) solution to the SDE*

$$dY_t = \mu(Y_t)\, dt + \sigma(Y_t)\, dZ_t, \quad Y_0 > 0. \tag{A.6}$$

Suppose, moreover, that Y^* is a process satisfying

$$Y_t^* = Y_0 + \int_0^t \eta_s\, ds + \int_0^t \sigma(Y_t^*)\, dZ_t, \tag{A.7}$$

where η is a progressively measurable process such that $\eta_t \geq \mu(Y_t)$ for all t. Then $Y_t^* \geq Y_t$ for all t almost surely.

PROOF: The proof shown in Ikeda and Watanabe (1981, pp. 168-170) implies existence and uniqueness of the solution to $(A.6)$. For the second assertion, we can extend a standard

SDE comparison result (for example, as in Protter (1990)). An extension is called for since the usual (Gronwall-inequality-based) proof relies on a Lipschitz condition for the diffusion. It is enough to show that $E[(Y_t - Y_t^*)^+] = 0$ for any time t, which we will do with a slight variation of the Ikeda-Watanabe uniquess proof. Let

$$\varphi_n(x) = \int_0^{x^+} \int_0^y \psi_n(u)\, du\, dy, \quad x \in I\!R,$$

where ψ_n is defined exactly as in Ikeda and Watanabe (1981), pages 168-169, in terms of the function ρ satisfying the properties specified in the Yamada Condition. Almost exactly as in Ikeda and Watanabe, we have $\varphi_n \in C^2(I\!R)$, $0 \leq \varphi_n'(u) \leq 1$, and $\varphi_n(u) \uparrow u^+$ as $n \to \infty$. Now,

$$\begin{aligned} E[\varphi_n(Y_t - Y_t^*)] &= E\left[\int_0^t \varphi_n'(Y_s - Y_s^*)[\mu(Y_s) - \eta_s]\, ds\right] \\ &\quad + \frac{1}{2}E\left[\int_0^t \varphi_n''(Y_s - Y_s^*)[\sigma(Y_s) - \sigma(Y_s^*)]^2\, ds\right] \\ &\leq \frac{t}{n}. \end{aligned}$$

The equality is an application of Ito's Lemma, using the fact that φ_n' and σ are bounded. The inequality follows from the negativity of the first expectation, the Yamada condition, and the fact that $\varphi''(u) \leq 2\rho^{-2}(|u|)/n$, following the calculations in Ikeda and Watanabe (1981, pp. 168-170). Letting n pass to infinity, dominated convergence implies that that

$$0 \leq E[(Y_t - Y_t^*)^+] = \lim_n E[\varphi_n(Y_t - Y_t^*)] \leq \lim_n \frac{t}{n} = 0,$$

which is the desired result. ∎

Proof of Theorem.

First we take the case in which $v_i(x) = v(x) = \alpha + \beta \cdot x$ for all i. Then we generalize.

Let $\{\epsilon_n\}$ be a positive strictly decreasing sequence of numbers converging to zero. For each n, let $X^{(n)}$ be the solution of the stochastic differential equation defined by (4.1) for $t \leq \tau_n = \inf\{s : v(X_s^{(n)}) = \epsilon_n\}$, and by $X^{(n)}(t) = X^{(n)}(\tau_n)$ for $t \geq \tau_n$. This is the process satisfying (4.1) that is absorbed at the boundary $\{x : v(x) = \epsilon_n\}$. Since the coefficient functions defining (4.1) are uniformly Lipschitz on the domain $\{x : v(x) \geq \epsilon_n\}$, $X^{(n)}$ is uniquely well defined and is a strong Markov process by standard SDE results.

With $\tau_0 = 0$, we can now define a unique process X on the closed time interval $[0, \infty]$ by $X_t = X_t^{(n)}$ for $\tau_{n-1} \leq t \leq \tau_n$, and by $X_t = x_0$ for $t \geq \tau \equiv \lim_n \tau_n$. If $\tau = +\infty$ almost surely, then X uniquely solves (4.1) on $[0, \infty)$, as desired, and is strong Markov.

We let $V_t = v(X_t)$, the "volatility" process, write $dV_t = \eta(X_t)\,dt + \sqrt{V_t}\beta^\top \Sigma\, dW_t$, where $\eta(x) = \beta^\top(ax+b)$. Without loss of generality, we can assume that ϵ_1 is close enough to 0 that, using Condition A, we have a constant $\bar{\eta} > \beta^\top \Sigma\Sigma^\top \beta/2$ such that $\eta(x) \geq \bar{\eta} > 0$ for all x in the strip $\{x : 0 \leq v(x) \leq \epsilon_1\}$. We can assume that $v(x_0) > \epsilon_1$, also without loss of generality. We construct below a strictly positive "comparison volatility" process $\hat{V}$ such that $V_t \geq \hat{V}_t$ for all t almost surely. With this, $\tau_n \geq \hat{\tau}_n = \inf\{t : \hat{V}_t = \epsilon_n\} \to +\infty$ almost surely, completing the proof.

In order to construct $\hat{V}$, we first construct the "excursions" of X defined by passages of $v(X_t)$ from ϵ_2 to ϵ_1. The excursion time intervals are $[T(i), T^*(i)]$, where $T^*(0) = 0$ and, for $i \geq 1$,

$$T(i) = \inf\{t \geq T^*(i-1) : v(X_t) = \epsilon_2\}; \quad T^*(i) = \inf\{t \geq T(i) : v(X_t) = \epsilon_1\}.$$

For $t \in [T(i), T^*(i)]$, let

$$\hat{V}_t = \epsilon_2 + (t - T(i))\bar{\eta} + \int_{T(i)}^{t} \sqrt{\hat{V}_t}\, dZ_t, \tag{A.8}$$

where $Z = \beta^\top \Sigma W$ (that is, Z is a multiple of a standard Brownian motion). For t in other (non-excursion) intervals $[T^*(i), T(i+1)]$, let $\hat{V}_t = V_t$. The process $\hat{V}$ is strictly positive. This is obvious off excursions, and during excursions follows from Ikeda and Watanabe (1981).

We claim that $\hat{V}_t \leq V_t$ for all t almost surely. Clearly this inequality is maintained off excursions. During the i-th excursion, V is given by

$$V_t = \epsilon_2 + \int_{T(i)}^{t} \eta(X_s)\, ds + \int_{T(i)}^{t} \sqrt{V_s}\, dZ_s. \tag{A.9}$$

Since $\bar{\eta} < \eta(X_t)$ for all $t \in [T(i), T^*(i)]$, the comparison Lemma appearing before this proof shows that $\hat{V}_t \leq V_t$ during excursions, almost surely. Thus $\hat{V}_t \leq V_t$ for all t, almost surely. The proof is complete in the case considered, of a single stochastic volatility factor.

Now, for the general case, let

$$D_n = \{x \in D : v_i(x) \geq \epsilon_n, i \in \{1, \ldots, n\}\}.$$

As before, there is a unique solution $X^{(n)}$ to (4.1) in D_n up to the hitting time $\tau_n = \inf\{t : \min_i v_i(X_t^{(n)}) = \epsilon_n\}$, and we let $X^{(n)}(t) = X^{(n)}(\tau_n)$ for $t \geq \tau_n$. Again we define X as the limit process. The proof proceeds as before, except that there is a volatility comparison process $\hat{V}_i$ for each i, defined i by i as above. By exploiting part (b) of Condition A, essentially the same arguments as above show that, for all i, we have $v_i(X_t) \geq \hat{V}_{it} > 0$ for all t almost surely. The result then follows as in the simpler case first considered.

References

W. Ames (1977) *Numerical Methods For Partial Differential Equations, 2nd edition.* New York: Academic Press.

J.-P. Ansel and C. Stricker (1991) "Lois de Martingale, Densités et Décompositions de Föllmer Schweizer," Université de Franche-Comté.

D. Beaglehole and M. Tenney (1991) "General Soltutions of Some Interest Rate Contingent Claim Pricing Equations," *Journal of Fixed Income* **1:** 69-83.

F. Black, E. Derman, and W. Toy (1990) "A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options," *Financial Analysts Journal*, January-February, pp. 33-39.

F. Black and M. Scholes (1973) "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* **81:** 637-654.

M. Brennan and E. Schwartz (1979) "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance* **3:** 133-155.

R. Brown and S. Schaefer (1993) "Interest Rate Volatility and the Term Structure of Interest Rates," *Philosophical Transactions of The Royal Society: Physical Sciences and Engineering* **347:** 563-576.

Y.-K. Chan (1992) "Term Structure as a Second Order Dynamical System and Pricing of Derivative Securities," Bear Stearns and Company.

L. Chen (1993) "Stochastic Mean and Stochastic Volatility: a Three-Factor Model of the Term Structure and its Application in Pricing of Interest Rate Derivatives," Working Paper, Federal Reserve Board, Washington D.C..

R.-R. Chen and L. Scott (1992a) "Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure," *Review of Financial Studies* **5:** 613-636.

R.-R. Chen and L. Scott (1992b) "Maximum Likelihood Estimation for a Multi-Factor Equilibrium Model of the Term Structure of Interest Rates," Working Paper, Rutgers University and University of Georgia.

R.-R. Chen and L. Scott (1993) "Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests from a State-Space Model Using a Kalman Filter," Working Paper, Rutgers University and University of Georgia.

G. Constantinides (1992) "A Theory of the Nominal Structure of Interest Rates," *Review of Financial Studies* **5:** 531-552.

J. Cox, J. Ingersoll, and S. Ross (1985a) "A Theory of The Term Structure of Interest Rates," *Econometrica* **53:** 385-408.

J. Cox, J. Ingersoll, and S. Ross (1985b) "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica* **53:** 363-384.

S. Das (1993) "Jump-Diffusion Processes and the Bond Markets," Working Paper, Harvard Business School, Revised, November, 1994.

M. Dothan (1978) "On the Term Structure of Interest Rates," *Journal of Financial Economics* **7:** 229-264.

D. Duffie (1992) *Dynamic Asset Pricing Theory.* Princeton University Press.

D. Duffie, J. Ma, and J. Yong (1993) "Black's Consol Rate Conjecture," Working Paper, Graduate School of Business, Stanford University, forthcoming: *Annals of Applied Probability.*

D. Duffie and K. Singleton (1995) "An Econometric Model of The Term Structure of Interest Rate Swap Yields," Working Paper, Graduate School of Business, Stanford University.

N. El Karoui and V. Lacoste (1992) "Multifactor Models of the Term Strucutre of Interest Rates," Working Paper, June, University of Paris VI.

N. El Karoui, R. Myneni, and R. Viswanathan (1992) "Arbitrage Pricing and Hedging of Interest Rate Claims with State Variables: I Theory," Working Paper, January, University of Paris VI.

N. El Karoui and J.-C. Rochet (1989) "A Pricing Formula for Options on Coupon Bonds," Working Paper, October, University of Paris VI.

A. Frachot, D. Janci, and V. Lacoste (1992) "Factor Analysis of the Term Structure: A Probabilistic Approach," Banque de France.

A. Frachot and J.-P. Lesne (1993) "Econometrics of Linear Factor Models of Interest Rates," Banque de France.

A. Friedman (1975) *Stochastic Differential Equations and Applications, Volume 1.* New York: Academic Press.

M. Harrison and D. Kreps (1979) "Martingales and Arbitrage in Multiperiod Security Markets," *Journal of Economic Theory* **20:** 381-408.

M. Harrison and S. Pliska (1981) "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Their Applications* **11:** 215-260.

D. Heath, R. Jarrow, and A. Morton (1992) "Bond Pricing and The Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica* **60:** 77-105.

S. Heston (1991) "Testing Continuous-Time Models of the Term Structure of Interest Rates," School of Organization and Management, Yale University.

F. E. Hohn (1964) *Elementary Matrix Algebra, second edition.* New York: Macmillan and Company.

T. Ho and S. Lee (1986) "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* **41:** 1011-1029.

C.-F. Huang (1987) "An Intertemporal General Equilibrium Asset Pricing Model: The Case of Diffusion Information," *Econometrica* **55:** 117-142.

J. Hull and A. White (1990) "Pricing Interest Rate Derivative Securities," *Review of Financial Studies* **3:** 573-592.

N. Ikeda and S. Watanabe (1981) *Stochastic Differential Equations and Diffusion Processes.* Amsterdam: North-Holland.

F. Jamshidian (1989) "An Exact Bond Option Formula," *Journal of Finance* **44:** 205-209.

F. Jamshidian (1991) "Bond and Option Evaluation in the Gaussian Interest Rate Model and Implementation," *Research in Finance* **9:** 131-170.

F. Jamshidian (1992) "A Simple Class of Square-Root Interest Rate Models," Working Paper, Fuji International, London.

F. Jamshidian (1993) "Bond, Futures and Option Evaluation in the Quadratic Interest Rate Model," Working Paper, Fuji International, London.

T. Langetieg (1980) "A Multivariate Model of the Term Structure of Interest Rates," *Journal of Finance* **35:** 71-97.

R. Litterman and J. Scheinkman (1988) "Common Factors Affecting Bond Returns," Research Paper, Goldman Sachs Finanical Strategies Group.

F. Longstaff and E. Schwartz (1992) "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," *Journal of Finance* **47:** 1259-1282.

M. Musiela (1994) "Stochastic PDEs and Term Structure Models," Working Paper, School of Mathematics, University of New South Wales, Australia.

L. T. Nielsen and J. Saá-Requejo (1992) "Exchange Rate and Term Strcuture Dynamics and the Pricing of Derivative Securities," INSEAD working paper.

N. Pearson and T.-S. Sun (1994) "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, Ross Model," *Journal of Finance* **49:** 1279-1304.

G. Pennachi (1991) "Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data," *Review of Financial Studies* **4:** 53-86.

W. Press, B. Flannery, S. Teukolsky and W. Vetterling (1988)) *Numerical Recipes in C.* Cambridge University Press.

P. Protter (1990) *Stochastic Integration and Differential Equations.* New York: Springer-Verlag.

S. Richard (1978) "An Arbitrage Model of the Term Structure of Interest Rates," *Journal of Financial Economics* **6:** 33-57.

C. Rogers (1995) "Which Model for the Term Structure of Interest Rates Should One Use?." **In** M. Davis, D. Duffie, W. Fleming, and S. Shreve, *Mathematical Finance.* New York: Springer Verlag, pp. 93-116.

R. Stambaugh (1988) "The Information in Forward Rates: Implications for Models of the Term Structure," *Journal of Financial Economics* **21:** 41-70.

O. Vasicek (1977) "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics* **5:** 177-188.