Non-Negative Affine Yield Models of the Term Structure

Kin Pang and Stewart D Hodges

Abstract

Affine yield models of the term structure (e.g. as described by Duffie & Kan, 1995) form an important general class which encompasses many popular models. However, it also includes models with the unfortunate property that interest rates can go negative. In practice, it would be difficult to estimate the coefficients of this form of model to preclude negative rates occurring. In this paper we show that those affine yield models which guarantee only positive interest rates are equivalent to generalised versions of the Cox, Ingersoll and Ross (1985) model. Working with the model in this framework provides an easier way to guarantee that all yields remain non-negative.

1 INTRODUCTION

Duffie & Kan (1995), DK, presents a consistent and arbitrage-free affine multi-factor model of the term structure of interest rates in which the factors can be chosen to be bond yields at selected maturities that follow a parametric multi-variate Markov diffusion process with stochastic volatility. We refer to their model with arbitrary factors as the Duffie and Kan Affine Model and their model with the factors specified as yields as the Duffie and Kan Affine Yield Model. The Duffie and Kan Affine Model generalises several other published affine models, which we call the simple affine models. The models are affine in the sense that bond yields are affine functions of the factors. Examples of simple affine models in the literature include Vasicek (1977), Langetieg (1980), Cox, Ingersoll & Ross (1985) , Longstaff & Schwartz (1992), Fong & Vasicek (1992) and Chen & Scott (1992).

For any model to be popular with practitioners, it must be both mathematically tractable and easy to calibrate to market data. Duffie and Kan argue that their affine yield models are more attractive than simple affine models. There are a number of reasons why this may be the case. Firstly, since the factors are yields which are observable and whose covariance structure can easily be measured or implied from market data, the calibration of the Duffie and Kan Affine Yield Models should be more transparent and easier than of the simple affine yield models. The simple affine models of Longstaff & Schwartz and Fong & Vasicek, for example, have yields and their volatilities as functions of the short rate and volatility of the short rate which cannot be observed[1]. Secondly, yields are direct inputs to the Duffie and Kan Affine Yield Model, so changing model parameters only changes the covariance structure. This property again aids calibration as changing some parameters of the simple affine models changes both the yields and their covariance structure making the numerical search for the optimal parameters more difficult. Thirdly, yields and their covariance structure are the natural variables to consider when pricing many derivatives such as caps and swaptions. Having yields as the factors not only makes the calibration simpler but also provides more intuition and insight into the valuation and hedging of interest rate derivatives.

The advantages of the Duffie and Kan Affine Yield Model are counterbalanced by the fact that the parameters of the Duffie and Kan yields process must satisfy more constraints than those of the simple affine models. Some of those constraints are interior and terminal conditions on a system of Ricatti equations. In general we have to employ numerical techniques to find suitable parameters that satisfy those conditions. This is a serious problem and finding appropriate parameter values may be even more demanding than the calibration issues of the simple affine factor models highlighted earlier. Finding parameters consistent with non-negative yields is even more difficult.

The difficulties just described for obtaining a Duffie and Kan Non-Negative Affine Yield Model lead us to believe that we may still therefore wish to work with the simple affine models as it is far easier to generate non-negative yields from them.

---

[1] Clewlow & Strickland (1994) examines the practical difficulties of estimating the Longstaff & Schwartz interest rate model from financial data.

Some practitioners may use models that allow negative yields when they offer compensating advantages of analytical tractability and computational speed. They sometimes argue they can choose model parameters to allow only a small probability of yields going negative. However Rogers (1993) shows that in the basic Vasicek (1977) interest rate model, even a small probability of negative spot rates can have a much larger effect on bond prices and hence on more complicated derivatives, especially when the maturities are distant. We would expect the same consequence for a Duffie and Kan Affine Yield Model that permits negative interest rates. We therefore focus our attention specifically on the Duffie and Kan Non-Negative Affine Yield Model.

The Duffie and Kan Non-Negative Affine Yield Model offers a large number of degrees of freedom and at first sight that flexibility appears to allow for models beyond that currently in the literature. However, this paper establishes that the Duffie and Kan Non-Negative Affine Yield Models are equivalent to a subset of the Duffie and Kan Affine Models, the Generalised Cox Ingersoll and Ross (CIR) Interest Rate Models, for which parameters can be easily chosen to ensure non-negative yields. This equivalence result provides an alternative method for calibrating Duffie and Kan Affine Yield Models consistent with non-negative yields.

We begin in Section 2 by showing how the Duffie and Kan class of models is related to models in the current literature, providing an introduction to the Duffie and Kan Affine Models and defining the Generalised CIR Interest Rate Models. Section 3 derives a number of results finishing with the equivalence between Duffie and Kan Non-Negative Affine Yield Models and Generalised CIR Interest Rate Models. Section 4 compares two approaches to calibrating a Duffie and Kan Non-Negative Affine Yield Model: An indirect method that calibrates a Generalised CIR Interest

Rate Model before converting it to its Duffie and Kan Non-Negative Affine Yield Model equivalent and a direct approach that works with the yields process and short rate specification. Section 5 summarises. We adopt the notation of DK throughout.

2 THE DUFFIE & KAN AFFINE MODELS AND THE GENERALISED COX INGERSOLL AND ROSS INTEREST RATE MODEL

In this section we first relate the class of Duffie & Kan Affine Models to the current literature before proceeding to describe the equations which characterise them. We also define a strict subset of the Duffie & Kan Affine Model we call the Generalised CIR Interest Rate Model. Readers are referred to DK for more details.

The Duffie and Kan Affine class of models offer some tractability and is a generalisation of many models in the literature. The Duffie and Kan Affine class includes as special cases the well known models of Vasicek (1977), Langetieg (1980), Cox, Ingersoll & Ross (1985), Fong & Vasicek (1991), Longstaff & Schwartz (1992) and Chen & Scott (1992). In fact, nearly all interest rate models with analytic bond pricing formulae belong to the Duffie and Kan Affine class. The Duffie and Kan parameters can also be allowed to be time dependent so that the Duffie and Kan Affine class also includes extended versions of the above models. Allowing the parameters to be time-dependent enables all of the above models to fit the entire term structure of interest rates. This practice of allowing for time dependence is sometimes called recalibration. The popular lognormal interest rate models of Black, Derman & Toy (1990) and Black & Karasinski (1991) however do not belong to the Duffie and Kan Affine class and do not have analytic bond pricing formulae. The three factor stochastic mean and stochastic volatility interest rate model of Chen (1994) and the lognormal model of Brace, Gaterek and Musiela (1995) are rare examples of non-Duffie and Kan Affine Models with analytic bond pricing formulae. All Duffie and Kan Models are special cases of Heath, Jarrow and Morton (1992): The evolution of the entire instantaneous forward rate curve is completely characterised by a finite number of Markovian factors. Karoui, Geman and Lacoste (1995) provide conditions for those factors to be yields when yield volatilities are deterministic in the Gaussian Heath Jarrow and Morton framework.

2A Duffie & Kan Affine Model

For the n factors $\underline{X}$, Duffie and Kan show that to achieve a bond pricing formula of the form

$$P(\underline{X},\tau) = \exp[A(\tau) + \underline{B}(\tau)^T \underline{X}] \tag{1}$$

where $A(\tau)$ is a function, $\underline{B}(\tau)$ is a column vector of n functions, τ is the bond maturity and P the bond price then the risk-neutral process for $\underline{X}$ takes the form[2]

$$d\underline{X} = (a\underline{X} + \underline{b})dt + \Sigma \begin{bmatrix} \sqrt{v_1(\underline{X})} & & 0 \\ & \ddots & \\ 0 & & \sqrt{v_n(\underline{X})} \end{bmatrix} d\underline{W} \tag{2}$$

with

$$v_i = \alpha_i + \underline{\beta}_i^T \underline{X} \tag{3a}$$

where a and Σ are (n x n) matrices, $\underline{b}$ and $\underline{\beta}_i$ are (n x 1) column vectors, $d\underline{W}$ is a (n x 1) column vector of independent Brownian increments, parameters are chosen to give $\underline{X}$ a unique solution[3] and the superscript T denotes matrix transpose. DK show that Σ has to be non-singular when none of the factors are scalings of each other. We assume this throughout. We define the matrix β as the matrix with vector β_i as its ith column so that we can write

$$\underline{v} = \underline{\alpha} + \beta^T \underline{X}. \tag{3b}$$

[2] Note condition given in proposition of page 9, DK, for this to be the case.

[3] Duffie & Kan, pp. 10-12.

Specifying the short rate, r, as an affine function of the factors $\underline{X}$,

$$r = f + \underline{G}^{\mathrm{T}}\underline{X} \tag{4}$$

completes the specification of the Duffie and Kan Affine Model. To provide a more compact notation we use

$$\underline{v}^D \equiv \begin{bmatrix} v_1 & & 0 \\ & \ddots & \\ 0 & & v_n \end{bmatrix} \text{ and } \sqrt{\underline{v}^D} \equiv \begin{bmatrix} \sqrt{v_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{v_n} \end{bmatrix}$$

henceforth. We use $\underline{u}_i$ to denote an unit column vector with zeros for all its elements except for the ith element which has the entry 1. DK show that for all i the conditions

C1) For all $\underline{x}$ such that $v_i(\underline{x}) = 0$, $\underline{u}_i^T \beta^T (a\underline{x}+\underline{b}) > \dfrac{\underline{u}_i^T \beta^T \Sigma\Sigma^T \beta \underline{u}_i}{2}$ and

C2) For all j, if $(\beta^{\mathrm{T}}\Sigma)_{ij} \neq 0$, then $v_i = v_j$

ensure that $\underline{v} > \underline{0}$. Intuitively, condition (C1) provides for a sufficiently large drift for $dv_i(\underline{x})$ when $v_i(\underline{x})$ approaches zero and condition (C2) prevents the correlation between $v_i(\underline{x})$ and $v_j(\underline{x})$ from taking $v_i(\underline{x})$ across zero.

Given the process for $\underline{X}$ and the specification of the short rate we can solve for the functions $A(\tau)$ and $B_i(\tau)$ for $i = 1$ to n, perhaps numerically, subject to the initial conditions

$$A(0) = 0 \text{ and } B_i(0) = 0. \tag{5}$$

For example, in a two factor model with $r = f + G_1X_1 + G_2X_2$ the functions $A(\tau)$ and $\underline{B}(\tau)$ are the solutions to the Ricatti equations

$$\frac{dB_1(\tau)}{d\tau} = -G_1 + a_{11}B_1(\tau) + a_{21}B_2(\tau) \tag{6a}$$
$$+\frac{\beta_{11}}{2}[\Sigma_{11}B_1(\tau) + \Sigma_{21}B_2(\tau)]^2 + \frac{\beta_{12}}{2}[\Sigma_{12}B_1(\tau) + \Sigma_{22}B_2(\tau)]^2$$

$$\frac{dB_2(\tau)}{d\tau} = -G_2 + a_{12}B_1(\tau) + a_{22}B_2(\tau) + \frac{\beta_{21}}{2}[\Sigma_{11}B_1(\tau) + \Sigma_{21}B_2(\tau)]^2 + \frac{\beta_{22}}{2}[\Sigma_{12}B_1(\tau) + \Sigma_{22}B_2(\tau)]^2 \tag{6b}$$

$$\frac{dA(\tau)}{d\tau} = -f + b_1B_1(\tau) + b_2B_2(\tau) + \frac{\alpha_1}{2}[\Sigma_{11}B_1(\tau) + \Sigma_{21}B_2(\tau)]^2 + \frac{\alpha_2}{2}[\Sigma_{12}B_1(\tau) + \Sigma_{22}B_2(\tau)]^2 \tag{6c}$$

subject to the initial conditions $A(0) = B_1(0) = B_2(0) = 0$. Note that the specification of the short rate directly affects the solution to the Ricatti equations. That is f and $\underline{G}$ are parameters of the model. The system of ordinary differential equations is a standard initial value problem which can be solved numerically using the fourth-order Runge-Kutta method as described in Press *et al* (1988).

2B The Duffie & Kan Affine Yield Model

When the underlying factors $\underline{X}$ are yields, in addition to providing a unique solution for the stochastic differential equation for $\underline{X}$, we must also choose parameters such that when we price a bond with a maturity τ_i that is the same as for one of the chosen yields, the bond pricing formula is given by $P(\underline{X}, \tau_i) = \exp[-\tau_i X_i]$. The parameters must therefore be chosen such that the bond pricing formula meets the following constraints:

$$A(\tau_j) = B_i(\tau_j) = 0,\ j \neq i,\ B_i(\tau_i) = -\tau_i \text{ for } i = 1 \text{ to } n. \tag{7}$$

The general bond pricing formula, equation (1), will then reduce to the required form when pricing a bond with a maturity that is the same as one of the chosen reference yields. As the functions $A(\tau)$ and $B_i(\tau)$ for $i = 1$ to n are solutions to Ricatti equations with interior and terminal conditions, finding suitable parameters to satisfy these conditions may be difficult. We may attempt to solve the system of ordinary differential equations by the shooting method as described in Press *et al* (1988).

2C Duffie & Kan Non-Negative Affine Yield Model

For bond yields to remain non-negative[4], we require the short rate be non-negative. Given that $A(\tau)$ and $\underline{B}(\tau)$ satisfy equations (5) and (7), Section 3B shows that the short rate is given by

$$r = \left[\left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{\alpha} - \left. \frac{\partial A(\tau)}{\partial \tau} \right|_{\tau=0} \right] - \left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{v} \tag{8}$$

so that sufficient conditions for non-negative yields are

$$\left[\left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{\alpha} - \left. \frac{\partial A(\tau)}{\partial \tau} \right|_{\tau=0} \right] \geq 0 \tag{9a}$$

and

$$- \left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \geq \underline{0}. \tag{9b}$$

---

[4] Yields in both the DK Non-Negative Affine Yield Model and the Generalised CIR Interest Rate Model can be made strictly positive by changing some of the conditions given for non-negative yields.

2D Converting a Duffie & Kan Affine Model to a Yield Model

DK show that we can transform the Duffie and Kan Affine Model $P(\underline{X},\tau) = \exp[A(\tau) + \underline{B}(\tau)^T \underline{X}]$ where $\underline{X}$ are general factors (which may or may not be yields) to the Duffie and Kan Affine Yield Model with yields of maturities τ_j, $j = 1$ to n by the substitution $\underline{Y} = \underline{g} + h^T\underline{X}$ whenever the matrix $h = -\left[\frac{B_i(\tau_j)}{\tau_j}\right]_{i,j}$ is non-singular.

The vector $\underline{g}$ is given by $-\left[\frac{A(\tau_j)}{\tau_j}\right]_j$.

PROPOSITION 1: *We can always find a set of distinct τ_j, $j = 1$ to n such that h is non-singular when the model is non-degenerate.*

By non-degenerate we mean that no variables are redundant in the sense that it is not possible to find a change of variables that allows the bond pricing formula to be specified with fewer factors. We show that if h is singular then the bond pricing formula has at least one redundancy.

PROOF: Suppose h is singular for distinct non-zero choices of τ_j, $j = 1$ to n.

Define

$$\mathrm{B} = \begin{bmatrix} B_1(\tau_1) & \cdots & B_1(\tau_n) \\ \vdots & & \vdots \\ B_n(\tau_1) & \cdots & B_n(\tau_n) \end{bmatrix}.$$

Then equation (1) implies

$$\begin{bmatrix} \frac{dP(\tau_1)}{P(\tau_1)} \\ \vdots \\ \frac{dP(\tau_n)}{P(\tau_n)} \end{bmatrix} = ()dt + \mathrm{B}^T d\underline{X} \tag{10}$$

where B is singular if and only if h is singular[5] when none of the chosen maturities is zero. As B is singular there exists a $\underline{\lambda}$ with at least two non-zero elements such that $B\underline{\lambda} = \underline{0}$ since none of the columns of B is $\underline{0}$ when all the chosen reference yields are greater than zero. Define C to be identity matrix with a column replaced by $\underline{\lambda}$ such that C is non-singular. Then defining $E = B^T C$, $\underline{Y} = C^{-1}\underline{X} + \underline{D}$ and substituting in equation (10) gives

$$\begin{bmatrix} \frac{dP(\tau_1)}{P(\tau_1)} \\ \vdots \\ \frac{dP(\tau_n)}{P(\tau_n)} \end{bmatrix} = ()dt + Ed\underline{Y}. \tag{11}$$

The columns of the matrix E in equation (11) are the same as those of B except one which is $\underline{0}$ by construction. We have provided a change of variable such that the bond prices are now only driven by n-1 factors or less and therefore that the original formulation must have had at least one redundancy. QED.

We therefore assume there are no redundancies in the bond-pricing formula throughout this paper and that h therefore is always non-singular for distinct choices of maturities. We have shown that can transform any Duffie and Kan Affine Model to a Duffie and Kan Affine Yield Model where the maturities are non-zero and distinct. We can also therefore convert to a set of distinct yields that includes the short rate as the short rate is an affine function of the other yields.

---

[5] We have $h = -\begin{bmatrix} B_1(\tau_1) & \cdots & B_1(\tau_n) \\ \vdots & & \vdots \\ B_n(\tau_1) & \cdots & B_n(\tau_n) \end{bmatrix} \begin{bmatrix} 1/\tau_1 & & 0 \\ & \ddots & \\ 0 & & 1/\tau_n \end{bmatrix}$.

2E The Generalized CIR Interest Rate Model

The Generalized CIR Interest Rate Model is the model that is obtained when we specify the short rate as an affine function of $\underline{Y}$ that follows the process of DK with the restriction that the matrix Σ be the identity matrix and $v_i(\underline{Y}) = Y_i$. That is

$$d\underline{Y} = (a\underline{Y} + \underline{b})\,dt + \sqrt{\underline{Y}^D}\,d\underline{W} \text{ and } r = f + \underline{G}^T\underline{Y}. \tag{12}$$

We say the process $d\underline{Y}$ has the CIR property. This is a generalisation of the 2-factor CIR models of Longstaff and Schwartz (1992) and Chen and Scott (1992). It allows for an arbitrary n factors with interactions between their drifts. To obtain the bond pricing formula we solve for the functions $A(\tau)$ and $B_i(\tau)$ for $i = 1$ to n as before. It is simple to guarantee that the yields are non-negative and we consider non-negative yields a defining property of the Generalised CIR Interest Rate Model. That is, all Generalised CIR Interest Rate Model have non-negative yields.

PROPOSITION 2: *The short rate is non-negative if*

i) *for all* $\underline{Y}$ *with* $y_i = 0$, $(a\underline{Y}+\underline{b})_i > 0$ *for* $i = 1$ *to* n;

ii) $\underline{G} \geq \underline{0}$ *and* $f \geq 0$.

PROOF: Condition (i) ensures non-negative $\underline{Y}$ and condition (ii) ensures that the short rate is a non-negative sum $\underline{Y}$ and a non-negative elements. QED.

Condition (i) is satisfied when the off-diagonals of a are non-negative and the elements of $\underline{b}$ are greater than zero. Zero is a reflecting barrier for the short rate and implies bond yields are strictly positive. $\underline{Y}$ mean reverts to $-a^{-1}\underline{b}$ when a is negative definite.

The Generalised CIR Interest Rate Model is clearly a special case of the Duffie and Kan Affine Yield Model. We will go on to show that Generalised CIR Interest Rate Models are equivalent to Duffie and Kan Non-Negative Affine Yield Models.

3 EQUIVALENCE BETWEEN GENERALIZED CIR INTEREST RATE AND DUFFIE & KAN NON-NEGATIVE AFFINE YIELD MODELS

As a preliminary step, we show that in the Duffie and Kan Non-Negative Affine Yield Model the matrix β must be non-singular. This allows us to count the degrees of freedom present in the Duffie and Kan Non-Negative Yield Model and observe that it is the same as that in the Generalised CIR Interest Rate Model suggesting the two may be equivalent. The non-singularity of β allows us to transform uniquely from the Duffie and Kan reference yields to variables with the CIR property to establish the equivalence.

PROPOSITION 3: *In the Duffie and Kan Non-Negative Affine Yield Model, the matrix β is non-singular.*

The intuition of the proof is as follows: The state space for the stochastic differential equations (2) and (3) is contained by the intersection of half spaces defined by $v_i \geq 0$ for $i = 1$ to n. For the state space to contain only non-negative $\underline{X}$, it must be the case that the state space is bounded by a number of non-parallel boundaries equal to or greater than the dimension of the hyperspace and this implies β must be non-singular.

PROOF: Consider the equation $\left[\underline{\beta}_1 \quad \cdots \quad \underline{\beta}_n\right]^T \Delta\underline{x} = \underline{0}$ which only has non-degenerate solutions $\Delta\underline{x} \neq \underline{0}$ when β is singular. $\Delta\underline{x}$ is a vector that is parallel to all the boundaries. Suppose β is singular and $\hat{\underline{X}}$ is a feasible point, $\underline{v}(\hat{\underline{X}}) > \underline{0}$. It follows that $\hat{\underline{X}} + k\Delta\underline{x}$ is also a feasible point since $d\underline{v} = k\beta^T\Delta\underline{x} = \underline{0}$. However $\hat{\underline{X}} + k\Delta\underline{x} \geq \underline{0}$ cannot hold for all k, that is not all yields can be non-negative when β is singular.

Hence β must be non-singular in the Duffie and Kan Non-Negative Affine Yield Model[6]. QED.

3A Degrees of Freedom

We first introduce a restriction to Duffie and Kan Affine Yield Model specification to remove some redundant parameters and then show that a Duffie and Kan Affine Yield Model has the same number of degrees of freedom as a Generalised CIR Interest Rate Model.

Consider equarions (2) and (3b) where both β and Σ are non-singular. DK shows that $\beta^T\Sigma$ must be a diagonal matrix as $v_i \neq v_j$ for $i \neq j$ when β is non-singular. The restriction is necessary for non-negative $\underline{v}$ and follows immediately from an examination of the process for $\underline{v}$[7].

PROPOSITION 4: *We can scale uniquely the matrices β and Σ such that $\beta^T\Sigma = I$ without making any difference to the stochastic process.*

PROOF: Define $\overline{\Sigma} = \Sigma D$, $\overline{\underline{v}} = D^{-2}\underline{\alpha} + D^{-2}\beta^T\underline{X} = \overline{\underline{\alpha}} + \overline{\beta}^T\underline{X}$. Then $d\underline{X} = (a\underline{X} + \underline{b})dt + \overline{\Sigma}\sqrt{\overline{v}}^D d\underline{W}$ with $\overline{\beta}^T\overline{\Sigma} = I$ and the covariance structure of $d\underline{X}$ is given by $\overline{\Sigma}\sqrt{\overline{\underline{v}}}^D\sqrt{\overline{\underline{v}}}^D\overline{\Sigma}^T = \overline{\Sigma}\overline{v}^D\overline{\Sigma}^T = \Sigma D\overline{\underline{v}}^D D^T\Sigma^T = \Sigma\underline{v}^D\Sigma^T$ which is the same as before. Thus the stochastic process for $d\underline{X}$ is the same and we can always choose β

---

[6] Alternatively, one can argue that none of the boundaries can be parallel for there to exist solutions to the SDE.

[7] $d\underline{v} = \left[\beta^T a(\beta^T)^{-1}\underline{v} + (\beta^T\underline{b} - \beta^T a(\beta^T)^{-1}\underline{\alpha})\right]dt + \beta^T\Sigma\sqrt{\underline{v}}^D d\underline{W}$.

and Σ such that $\beta^T \Sigma = I$ in the Duffie and Kan Affine Non-Negative Yield Model.

QED.

PROPOSITION 5: *The Duffie and Kan Non-Negative Affine Yield Model has the same degrees of freedom as the Generalised Cox Ingersoll and Ross Interest Rate Model.*

PROOF: Assume the chosen yields do not include the short rate. There are $2n^2 + 2n$ parameters in the yields process. There are n+1 parameters for the specification of the short rate. The bond pricing formula, equation (1), must satisfy the following restrictions:

	τ_1	----	τ_i	----	τ_n
$A(\tau)$	$A(\tau_1) = 0$	----	$A(\tau_i) = 0$	----	$A(\tau_n) = 0$
$B_1(\tau)$	$B_1(\tau_1) = -\tau_1$	----	$B_1(\tau_i) = 0$	----	$B_n(\tau_1) = 0$
:	:		:		:
$B_j(\tau)$	:		$B_j(\tau_i) = 0$ if $j \neq i$ $B_j(\tau_j) = -\tau_j$		:
:			:		:
$B_n(\tau)$	$B_n(\tau_1) = 0$	----	$B_n(\tau_i) = 0$	----	$B_n(\tau_n) = -\tau_n$

These give a total of $n(n + 1)$ terminal boundary conditions on the set of Ricatti equations for $A(\tau)$ and $B_j(\tau)$ so that the number of independent parameters in the Duffie and Kan Affine Yield Model is $(n+1)^2$.

The generalized CIR Interest Rate Model, equation (12) also has $(n+1)^2$ parameters, although we usually restrict f to be zero. Thus the Duffie and Kan Affine

Yield Model has the same degrees of freedom as the Generalised CIR Interest Rate Model[8]. QED.

To show that the Generalized CIR Interest Rate Model and the Duffie and Kan Non-Negative Affine Yield Model are equivalent we only need to show a Duffie and Kan Non-Negative Affine Yield Model can be transformed uniquely to a Generalised CIR Interest Rate Model since Section 2C has already shown that we can convert uniquely from an Duffie and Kan Affine Model to a Duffie and Kan Affine Yield Model.

3B Equivalence

PROPOSITION 6: *The Duffie and Kan Non-Negative Affine Yield Model and the Generalised Cox Ingersoll and Ross Interest Rate Models are equivalent.*

PROOF: Starting from a Duffie and Kan Non-Negative Affine Yield Model, we can express the short rate as a sum of CIR variables using the transformation defined by equation (3b). Eliminating $\underline{X}$ from the bond pricing formula, equation (1), gives equation (8) of section 2C

$$r = \left[\left.\frac{\partial \underline{B}(\tau)^T}{\partial \tau}\right|_{\tau=0} (\beta^T)^{-1} \underline{\alpha} - \left.\frac{\partial A(\tau)}{\partial \tau}\right|_{\tau=0} \right] - \left.\frac{\partial \underline{B}(\tau)^T}{\partial \tau}\right|_{\tau=0} (\beta^T)^{-1} \underline{v}$$

with

[8] For the case when the short rate is included as one of the reference yields, indicated by τ_1 say, the restrictions on $A(\tau_1)$ and $\underline{B}(\tau_1)$ are redundant because these restrictions are now the same as the initial conditions for maturity equal to zero and so there are n+1 fewer restrictions. There are also n+1 fewer degrees of freedom as $r = f + \underline{G}^T\underline{X}$ becomes $r = X_1$ so the total degrees of freedom is unaltered.

$$
\begin{aligned}
d\underline{v} &= (\beta^T a\underline{X} + \beta^T \underline{b})dt + \beta^T \Sigma\sqrt{\underline{v}^D}\, d\underline{W} \\
&= [\beta^T a(\beta^T)^{-1}\underline{v} + (\beta^T \underline{b} - \beta^T a(\beta^T)^{-1}\underline{\alpha})]dt + \beta^T \Sigma\sqrt{\underline{v}^D}\, d\underline{W} \\
&= (E\underline{v} + \underline{F})dt + \sqrt{\underline{v}^D}\, d\underline{W}
\end{aligned}
$$

and obvious substitutions. Section 3B has already shown that β and Σ can be scaled to make $\beta^T\Sigma$ equal to the identity matrix. The two equations define a Generalised CIR Interest Rate Model. Furthermore, this transformation must be unique up to a permutation of the indices. Any other affine transformation would not maintain the CIR property. We can therefore conclude that the two models are equivalent as Section 2D has already shown that we can transform uniquely in the opposite direction. QED.

We are now at the stage where we know the models are equivalent. We next look at the difficulties encountered in empirical work when working with the two different forms.

4 IMPLEMENTATION AND CALIBRATION ISSUES

We distinguish between implementation and calibration issues. Implementation problems are the problems of how to get a model to value instruments and calibration problems are those we face when we want the model prices to match market data.

Suppose we want a Duffie and Kan Non-Negative Affine Yield Model. In this section we consider two approaches for obtaining suitable parameters: an indirect method that first estimates a Generalised CIR Interest Rate Model and then transforms to its Duffie and Kan equivalent and a direct approach that estimates the parameters for the yields process. We first examine an indirect approach and find that the implementation is simple but the calibration is difficult because it is necessary to search over a large number of variables. We then look at a direct approach and find considerable implementation problems. It is difficult to choose parameters that are consistent with the factors being yields. Calibration would be relatively simple if there were a quick way to choose suitable consistent parameters. In both cases we assume we only have accurate estimates of the covariance structure between yields. We are estimating the risk-neutral processes and cannot therefore expect to have good estimates of the drift parameters from examining the objective behaviour of the yields. We do not use options prices as this is only a preliminary investigation.

4A The Indirect Approach

We first illustrate the ideas with the Longstaff & Schwartz two factor affine model before explaining how we may proceed when we do not have an explicit bond pricing formula. Both cases can be easily implemented to be consistent with non-negative yields for the constraints guaranteeing non-negativity are simple non-negativity and strict positivity constraints on model parameters made clear in Proposition 2 of Section 2E. Calibrating the models in both cases is more difficult.

Starting essentially from a specification of 2 CIR variables in their general equilibrium model

$$dx = (\gamma - \delta x)dt + \sqrt{x}dW_1 \text{ and } dy = (\eta - \xi y)dt + \sqrt{y}dW_2$$

Longstaff and Schwartz show that *r*, the short rate, and *V*, the instantaneous variance of the short rate, are given by

$$r = \alpha x + \beta y \text{ and } V = \alpha^2 x + \beta^2 y$$

where α and β are positive parameters of their model. They solve initially for a bond pricing formula in terms of *x* and *y* which they then through a change of variables express in terms of *r* and *V*. Longstaff & Schwartz have performed a change of factors rather like DK. Expressing the formula as a function of two different yields instead would give a Duffie and Kan Non-Negative Affine Yield Model.

Suppose we wish to establish a model with the instantaneous rate and the one year yield as the factors and we want to fit the model to the market yields and covariance structure between the two chosen yields. Transforming from the factors (*r*, *V*) to the short rate, *Y*(0), and the 1 year rate, *Y*(1), gives

$$P(\underline{Y},\tau) = \exp[E(\tau) + \underline{F}(\tau)^T \underline{Y}] \text{ where}$$

$$\underline{Y} = \begin{bmatrix} Y(0) \\ Y(1) \end{bmatrix},\ Y(0) \equiv r,\ \underline{F}(\tau) = \begin{bmatrix} F_0(\tau) \\ F_1(\tau) \end{bmatrix}$$

$$E(\tau) = 2\gamma \ln A(\tau) + 2\eta \ln B(\tau) + \kappa\tau - \frac{D(\tau)}{D(1)}[\kappa + 2\gamma \ln A(1) + 2\eta \ln B(1)]$$

$$F_0(\tau) = C(\tau) - \frac{D(\tau)C(1)}{D(1)},\quad F_1(\tau) = -\frac{D(\tau)}{D(1)}.$$

where $A(\tau)$, $B(\tau)$, $C(\tau)$ and $D(\tau)$ are the functions of Longstaff and Schwartz reproduced in Appendix 7A and $A(1)$, $B(1)$, $C(1)$ and $D(1)$ are the respective evaluated at $\tau = 1$. The yields process can be shown to be

$$d\underline{Y} = (R\underline{Y} + \underline{S})dt + T\begin{pmatrix} \sqrt{v_0} & 0 \\ 0 & \sqrt{v_1} \end{pmatrix} d\underline{W} \text{ with}$$

$$R = \begin{bmatrix} R_{00} & R_{01} \\ R_{10} & R_{11} \end{bmatrix}$$ where R_{00}, R_{01}, R_{10} and R_{11} are defined in Appendix 7B.

$$S = \begin{bmatrix} S_0 \\ S_1 \end{bmatrix}$$ where S_0 and S_1 are defined in Appendix 7B.

$$T = \begin{bmatrix} \alpha & \beta \\ -\alpha[C(1) + \alpha D(1)] & -\beta[C(1) + \beta D(1)] \end{bmatrix} \text{ and}$$

$$\underline{v} = \underline{g} + h^T \underline{Y} \text{ with}$$

$$\underline{g} = \begin{bmatrix} g_0 \\ g_1 \end{bmatrix}$$ where h^T, g_0 and g_1 are defined in Appendix 7B.

In this case, all model properties are explicit functions of the model parameters. We calibrate the model to best fit five targets, the short rate, one year rate, and their covariance structure, using six variables α, β, γ, δ, η, and ν where $\nu = \xi + \lambda$ and λ is a risk parameter. We could minimise a weighted sum of squared differences between model and market values using a numerical search routine. Note that it may not be possible to obtain an exact fit with the five target values even though we are searching over six parameters. For instance, it is difficult to obtain a small correlation between yields in Longstaff and Schwartz's model. Note also that we do not need to be able to observe x and y, the two CIR variables. Their values are implicitly defined by the yields we are fitting the model to.

For Generalised CIR Interest Rate Models where we do not have analytic bond pricing formulae we can easily solve numerically equation (6) which in this case simplifies to

$$\frac{dB_1(\tau)}{d\tau} = -G_1 + a_{11}B_1(\tau) + a_{21}B_2(\tau) + \frac{1}{2}B_1(\tau)^2 \tag{13a}$$

$$\frac{dB_2(\tau)}{d\tau} = -G_2 + a_{12}B_1(\tau) + a_{22}B_2(\tau) + \frac{1}{2}B_2(\tau)^2 \tag{13b}$$

$$\frac{dA(\tau)}{d\tau} = -f + b_1B_1(\tau) + b_2B_2(\tau) \tag{13c}$$

subject to $A(0) = 0$ and $\underline{B}(0) = \underline{0}$ and the conditions of Proposition 2. Solving the ODE (13) gives equation (1), where $\underline{X}$ are unobserved CIR variables, which can be transformed to a formula in terms of bond yields $\underline{Y}$ giving

$$P(\underline{Y},\tau) = \exp[C(\tau) + \underline{D}(\tau)^T \underline{Y}]$$

where

$$\underline{Y} = g + h^{\mathrm{T}}\underline{X}\,,\ h = -\left[\frac{B_i(\tau_j)}{\tau_j}\right]_{i,j} \text{ and } g = -\left[\frac{A(\tau_j)}{\tau_j}\right]_j .$$

The unobserved CIR variables are given by $\underline{X} = (h^{\mathrm{T}})^{-1}(\underline{Y}-g)$. We re-iterate the search routine, with each iteration solving the ordinary differential equations once, until we have a good fit to observed yields and their covariance structure. Note that changing the model parameters changes only the covariance structure and not the reference yields. The reference yields are inputs and changing the model parameters changes the $A(\tau)$ and $\underline{B}(\tau)$ functions: The unobserved variables are constrained to take on different values that give the same observed reference yields which may include the short rate. This property can make it easier to search for the best parameters.

We have shown that it is in principle possible to calibrate whether we have or have not got closed form solutions to the Ricatti equations. There is, however, a practical problem in that it is necessary to search over a large number of parameters which increases rapidly with the number of factors. There are $(n + 1)^2$ parameters in a n-factor model.

We have explained how we can calibrate all Duffie and Kan Non-Negative Affine Yield Models indirectly by our equivalence result. The next section shows that it is much harder to implement Duffie and Kan Non-Negative Affine Yield Models directly.

4B The Direct Approach

In this section we explain how we might calibrate Duffie and Kan Non-Negative Affine Yield Models directly. It is far more difficult to implement Duffie and Kan Affine Yield Models directly than by the indirect method just explained. We illustrate the difficulties with a two factor example. More factors will increase the difficulties still further.

Suppose we are given two yields of different maturities, their volatilities and correlation that we wish to calibrate the model to. We need to determine the parameters for equations (2), (3) and (4) such that $\underline{v} \geq \underline{0}$ and that the functions $A(\tau)$, $B_1(\tau)$ and $B_2(\tau)$ of equation (1) satisfy equations (5) and (7). The difficulties arise almost entirely from having to solve equation (6) subject to equation (7), the interior and terminal conditions of equation (6). Unless there are analytic solutions to this system of Ricatti equations, we shoot the functions forward, that is, we integrate the equations from τ equal to zero to the longest chosen reference maturity. Having completed the numerical integration we measure the discrepancies between the boundary conditions and the solution values. The discrepancy measure can be a weighted sum of squared differences. We must iterate the shooting to reduce the discrepancy measure to zero and only after the measure has been reduced to an acceptably small value do we calculate the covariance structure. Otherwise the bond pricing formula would not be consistent with having yields as the factors. However, the covariance structure produced may not match the observed structure well enough and we must then resolve the Ricatti equations for different parameters until we produce a close match to the observed covariance structure. A simple calibration procedure may consists of two nested iterations. The inner loop ensures that the functions satisfy the boundary conditions. The outer loop minimizes the discrepancy between the model and target covariance structures. The procedure is clearly difficult. Imposing the constraints of equation (9) to preclude negative yields makes the numerical search for consistent parameters even more demanding. We cannot neglect the extra constraints as the following example illustrates. The following parameters for the joint process followed by the 1 year and 12 year yields

$$A = \begin{bmatrix} -1.31193 & 3.55331 \\ -0.088328 & -0.138073 \end{bmatrix}, \quad B = \begin{bmatrix} -0.20058 \\ 0.024806 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} -0.038705 & 0.372175 \\ -0.008299 & 0.043145 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 23.0104 \\ 2.27766 \end{bmatrix},$$

$$\beta = \begin{bmatrix} 30.4077 & 5.84922 \\ -262.299 & -27.2783 \end{bmatrix}$$

with

$$r = 0.21608 + 1.98915X_1 - 3.25205X_2$$

give absolute one year and twelve year yield volatilities of 5.35% and 1.43% respectively and a correlation of 0.975 when the one year yield and twelve year yields are 2.8% and 6.03% respectively. These parameters however permit negative yields for when we convert to the Generalised CIR setup we obtain

$$r = 0.6y - 0.05x \text{ with}$$

$$dx = (0.05 - 0.55x)dt + \sqrt{x}dW_1,$$

$$dy = (0.2 - 0.9y)dt + \sqrt{y}dW_2 \text{ where } dW_1 \text{ and } dW_2 \text{ are independent}$$

$$x = 1.107692 \text{ and } y = 0.158974.$$

Since y can attain zero independently of x, the formula for the short rate shows clearly that the short rate and hence all other yields can attain negative values. This example demonstrates a drawback of the Duffie and Kan Affine Yield Model; it is difficult to understand how the processes are behaving from a quick examination of the parameters.

Our example has shown that it is usually rather difficult to implement a Duffie and Kan Non-Negative Affine Yield Model directly. The previous section demonstrated that we are in principle able to calibrate Generalised CIR Interest Rate Models with non-negative interest rates readily but there is the difficulty of searching over many parameters. Both the direct and indirect calibration methods are challenging tasks. It may be possible to combine the two approaches to calibrate a Duffie and Kan Non-negative Affine Yield Model. For example, if we could measure parameters defining the yields covariance structure there would be fewer parameters to search over. That possibility is left for a future paper.

5 SUMMARY

We have shown that Duffie and Kan Non-Negative Affine Yield Models are equivalent to the Generalised Cox-Ingersoll-Ross Interest Rate Model. We have looked at the implementation and calibration issues and argued that, when we want non-negative yield models, both the Generalised Cox Ingersoll and Ross Interest Rate Model and the Duffie and Kan Non-Negative Affine Yield Models are difficult to work with empirically.

6 APPENDIX

6A The Longstaff & Schwartz Bond Pricing Formula

Longstaff and Schwartz show that bond prices within their model economy, where r is the short rate and V is the instantaneous variance of the short rate, are given by

$$P(r,V,\tau) = A^{2\gamma}(\tau)B^{2\gamma}(\tau)\exp[\kappa\tau + C(\tau)r + D(\tau)V]$$

where

$$A(\tau) = \frac{2\Phi}{(\delta+\Phi)(\exp(\Phi\tau)-1)+2\Phi},$$

$$B(\tau) = \frac{2\Psi}{(\upsilon+\Psi)(\exp(\Psi\tau)-1)+2\Psi},$$

$$C(\tau) = \frac{\alpha\Phi(\exp(\Psi\tau)-1)B(\tau) - \beta\Psi(\exp(\Phi\tau)-1)A(\tau)}{\Phi\Psi(\beta-\alpha)},$$

$$D(\tau) = \frac{\Psi(\exp(\Phi\tau)-1)A(\tau) - \Phi(\exp(\Psi\tau)-1)B(\tau)}{\Phi\Psi(\beta-\alpha)},$$

and $\upsilon = \xi + \lambda$, $\Phi = \sqrt{2\alpha+\delta^2}$, $\Psi = \sqrt{2\beta+\upsilon^2}$, $\kappa = \gamma(\delta+\Phi)+\eta(\upsilon+\Psi)$.

6B Duffie and Kan Formulation Of Longstaff & Schwartz Using The Short Rate And The One Year Yield As The Factors

We show here the joint process for short rate, $Y(0)$, and one year yield, $Y(1)$, within the Longstaff & Schwartz model. The process is obtained by expressing $Y(1)$ as a function of V followed by substitution. We obtain

$$d\underline{Y} = (R\underline{Y} + \underline{S})dt + T\begin{pmatrix} \sqrt{v_0} & 0 \\ 0 & \sqrt{v_1} \end{pmatrix} d\underline{W} \quad \text{with } \underline{v} = \underline{g} + h^T \underline{Y}$$

where

$$R_{00} = -\frac{D(1)(\beta\delta - \alpha\upsilon) - (\upsilon - \delta)C(1)}{(\beta - \alpha)D(1)}, \quad R_{01} = \frac{(\upsilon - \delta)}{(\beta - \alpha)D(1)},$$

$$R_{10} = -C(1)\left[-\frac{\beta\delta - \alpha\upsilon}{\beta - \alpha} + \frac{(\upsilon - \delta)C(1)}{(\beta - \alpha)D(1)}\right] - D(1)\left[-\frac{\alpha\beta(\delta - \upsilon)}{\beta - \alpha} + \frac{(\beta\upsilon - \alpha\delta)C(1)}{(\beta - \alpha)D(1)}\right],$$

$$R_{11} = -\frac{(\upsilon - \delta)C(1)}{(\beta - \alpha)D(1)} - \frac{\beta\upsilon - \alpha\delta}{\beta - \alpha}$$

$$S_0 = \alpha\gamma + \beta\eta - \frac{(\upsilon - \delta)(-\kappa - 2\gamma \ln A(1) - 2\eta \ln B(1)}{(\beta - \alpha)D(1)}$$

$$S_1 = -C(1)\left[\alpha\gamma + \beta\eta - \frac{(\upsilon - \delta)(-\kappa - 2\gamma \ln A(1) - 2\eta \ln B(1))}{(\beta - \alpha)D(1)}\right] - D(1)\left[\alpha^2\gamma + \beta^2\eta - \frac{(\beta\upsilon - \alpha\delta)(-\kappa - 2\gamma \ln A(1) - 2\eta \ln B(1))}{(\beta - \alpha)D(1)}\right]$$

$$g_0 = \frac{\kappa + 2\gamma \ln A(1) + 2\eta \ln B(1))}{\alpha(\beta - \alpha)D(1)}, \quad g_1 = -\frac{\kappa + 2\gamma \ln A(1) + 2\eta \ln B(1))}{\beta(\beta - \alpha)D(1)}$$

$$h = \begin{bmatrix} \underline{h}_0 & \underline{h}_1 \end{bmatrix}, \quad \underline{h}_0 = \begin{bmatrix} \dfrac{C(1) + \beta D(1)}{\alpha(\beta - \alpha)D(1)} \\ \dfrac{1}{\alpha(\beta - \alpha)D(1)} \end{bmatrix}, \quad \underline{h}_1 = \begin{bmatrix} -\dfrac{C(1) + \alpha D(1)}{\beta(\beta - \alpha)D(1)} \\ -\dfrac{1}{\beta(\beta - \alpha)D(1)} \end{bmatrix}.$$

7 REFERENCES

1. Black F, Derman E, Toy W, "*A One-Factor Model of Interest rates and Its Application to Treasury Bond Options*", Financial Analysts Journal, Jan-Feb 1990, pp 33-39.
2. Black F, Karasinski P, "*Bond and Option Pricing When Short Rates are Lognormal*", Financial Analysts Journal, July-Aug 1991, pp 52-59.
3. Brace A, Gatarek D, Musiela M, "*The Market Model of Interest Rate Dynamics*", Working Paper, The University of NSW, May 1995.
4. Chen L, "*Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in Pricing in Interest Rate Derivatives*", Working Paper, Federal Reserve Board, October 1994.
5. Chen RR, Scott L, "*Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure.*", The Review of Financial Studies, Vol 5 1992, pp 613-636.
6. Clewlow LJ, Strickland CR, "*A Note on Parameter Estimation in the Two Factor Longstaff and Schwartz Interest Rate Model*", Journal of Fixed Income, March 1994, pp 95-100.
7. Cox JC, Ingersoll JE, Ross SA, "*A Theory of the Term Structure of Interest Rates.*", Econometrica, March 1985, pp 385-407.
8. Duffie D, Kan R, "*A Yield-Factor Model of Interest Rates.*", Working Paper, Graduate School of Business, Stanford University, January 1995.
9. El Karoui N, Geman H, Lacoste V, "*On the Role of State Variables in Interest Rate Models*", Working Paper, May 1995.

10. Fong HG, Vasicek OA, "*Interest rate Volatility as a Stochastic Factor.*", Working Paper, Gifford Fong Associates, February 1991.

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

12. Langetieg TC, "*A Multivariate Model of the Term Structure.*", Journal of Finance, March 1980, pp 71-97.

13. Longstaff FA, Schwartz ED, "*Interest-Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model.*", Journal of Finance, September 1992, pp 1259-1282.

14. Press, Teukolsky, Vetterlin, Flannery, "*Numerical Recipes in C*", Second Edition, Cambridge University Press.

15. Rogers LCG, "*Which Model of the Term Structure of Interest rates Should One Use?*" Proceedings of the IMA Workshop on Mathematical Finance, 1993, presented at the Annual Conference of the Financial Options Research Centre, University of Warwick, UK, 1994.

16. Vasicek OA, "*An Equilibrium Characterisation of the Term Structure.*", Journal of Financial Economics, Vol 5 1977, pp 177-188.