Taking Positive Interest Rates Seriously*

ENLIN PAN†

Consultant

LIUREN WU‡

Associate Professor of Finance, Zicklin School of Business, Baruch College

This draft: October 19, 2003

First draft: February 21, 2001

*We thank Yacine Ait-Sahalia, Peter Carr, and Massoud Heidari for insightful comments. All remaining errors are ours.

†123 N. Waukegan Road, Suite 207, Lake Bluff, IL 60044; tel: (847) 863-4747; enlin@earthlink.net.

‡One Bernard Baruch Way, Box B10-225, New York, NY 10010; tel: (646) 312-3509; fax: (646) 312-3451; Liuren_Wu@baruch.cuny.edu; http://faculty.baruch.cuny.edu/lwu/.

Taking Positive Interest Rates Seriously

ABSTRACT

We present a dynamic term structure model in which interest rates of all maturities are bounded from below at zero. Positivity and continuity, combined with no arbitrage, result in only one functional form for the term structure with three sources of risk. One dynamic factor controls the level of the interest rate and follows a special two-parameter square-root process under the risk-neutral measure. The two parameters of the process determine the other two sources of risk and act as two static factors. This model has no other parameters to estimate and hence bears no other risks.

JEL Classification Codes: E43, G12, G13.

Keywords: term structure; consistency; positivity; quadratic forms.

Taking Positive Interest Rates Seriously

Many term structure models have been proposed during the last two decades, yet most of these models imply positive probabilities of negative interest rates. Other models guarantee interest rate positivity, but very often imply that interest rates at certain maturities cannot go below a certain positive number. Asserting that an interest rate can be negative or cannot be lower than, say, three percent, is equally counterintuitive for academics and troublesome for practitioners. In this paper, we propose a dynamic term structure model where interest rates of all maturities are bounded below at exactly zero.

Such a reasonable and seemingly innocuous contention, together with the assumption of continuity and no arbitrage, generates several striking results. First, the term structure of interest rates collapses to one functional form, determined by the solution to a scalar Riccati equation. Second, the term structure is governed by exactly three sources of risk, only one of which is dynamic. This dynamic risk factor follows a special two-parameter square-root process under the risk-neutral measure, and the two parameters of the process determine the other two sources of risk. The model has no extra parameters in addition to these three risk factors.

The most surprising result is the collapse of dimensionality. We obtain the three sources of risk without any *a priori* assumption on the exact dimensionality of the state space. The dynamic factor controls the level of the interest rate curve. The two parameters control the slope and curvature of the yield curve. Although the two parameters can be time varying, their dynamics do not affect the pricing of the interest rates. Therefore, we regard them as static factors.

Despite its simplicity, our model captures the observed yield curve very well. In particular, the model captures nicely the well-documented hump shape in the term structure of forward rates. By a simple transformation, we can represent the whole term structure by the maximum forward rate, the maturity of the maximum forward rate, and the curvature of the forward rate curve at the maximum. We can also use the instantaneous interest rate (level), the slope, and the curvature of the forward-rate curve at the short end as the three factors. Such transformations not only comply with the empirical findings and intuition, but also simplify the daily fitting of the forward-rate curve.

To investigate the empirical performance of the model in fitting the term structure of interest rates, we calibrate the model to the weekly data of both U.S. Treasury yields and U.S. dollar swap rates over the eight years from December 14, 1994 to December 28, 2000. The model fits both markets well. The pricing errors are mostly within a few basis points. The estimation also generates a time series of the three factors from both markets. The intuitive explanation of the three factors further enhances our understanding of the two interest rate markets. We find that although the average level spread between the swap rates and the Treasury yields are small, the spread can become exceptionally large during credit events such as the late 1998 hedge fund crisis and during the Treasury liquidity squeeze in 2000.

The paper is organized as follows. Section I describes the relevant literature that forms the background of our study. Section II elaborates on how the contention of interest rate positivity and continuity collapses the dimensionality of the state space to three. In Section III, we analyze the properties and different representations of the three sources of risk. In Section IV, we fit the model to both the U.S. Treasury yields and the U.S. dollar swap rates. In Section V, we explore the possibility of adding jumps to such a model while maintaining positive interest rates. Section VI concludes.

I. Background

Many term structure models allow positive probabilities of negative interest rates. The inconsistency in terms of negative interest rates in these models is often excused on the ground of "good" empirical performance and "small probability" of negative interest rates. Although this is true in many cases, the values of some derivatives are extremely sensitive to the possibility of negative rates (Rogers (1996)). For such derivatives, the prices inferred from these "negative" interest rate models can be absurd.

The literature has taken three approaches in generating positive interest rates. The first approach specifies the instantaneous interest rate as a general quadratic function of some Gaussian state variable. Examples of quadratic term structure models include Ahn, Dittmar, and Gallant (2002), Beaglehole and Tenney (1991, 1992), Brandt and Chapman (2002), Brandt and Yaron (2001), Constantinides (1992), El Karoui, Myneni, and Viswanathan (1992), Jamshidian (1996), Leippold and Wu (2002, 2003), Longstaff (1989), and Rogers (1997). This approach can guarantee the positivity of the instantaneous interest rate by one parametric restriction. However, the underlying dynamics very often imply that interest rates at some other maturities can either become negative or cannot go below a certain positive number. Asserting that an interest rate can be negative or cannot be lower than, say, three percent, is equally absurd. For example, no rational traders are willing to offer free floors at any strictly positive level of interest rates. Our model is mostly related to this approach. Instead of assuming a quadratic form for only the instantaneous interest rate, we require that interest rates at all maturities are quadratic functions of a finite-dimensional state vector. We further constrain the functions to have no linear or constant terms so that all interest rates are bounded from below at exactly zero.

The second approach derives positive interest rates based on the specifications of the pricing kernel. For example, Flesaker and Hughston (1996) derive a condition on the discount bond price that guarantees positive interest rates. However, the rational log-normal model they come up with from this condition has several issues: The short rate implied from the model is bounded from both above and below, and the model remains arbitrage-free only up to a certain point (Babbs (1997)). Jin and Glasserman (2001) show how the framework of Heath, Jarrow, and Morton (1992) is related to the positive rate framework of Flesaker and Hughston (1996).

The third approach treats nominal interest rates as options and hence guarantee positive interest rates. Examples include Black (1995), Gorovoi and Linetsky (2003), and Rogers (1995). In addition, Goldstein and Keirstead (1997) generate positive interest rates by modeling them as processes with reflecting or absorbing boundaries at zero. However, these models are rarely analytically tractable.

The collapse of dimensionality to three under our model is consistent with the empirical findings of factor analysis in, among others, Litterman and Scheinkman (1991), Knez, Litterman, and Scheinkman (1994), and Heidari and Wu (2003). The dimension of three has also become the consensus choice in recent empirical works on model designs, e.g., Backus, Foresi, Mozumdar, and Wu (2001), Balduzzi, Das, Foresi, and Sundaram (1996), Chen and Scott (1993), Dai and Singleton (2000, 2002, 2003), and Duffee (2002). However, these three-factor models have ten to 20 free parameters. The estimates of many of these parameters show large standard errors. Therefore, in applying these models, we not only need to control and price the risk of the three state variables (factors), but we must also be concerned with the uncertainty and risk associated with the many parameter estimates. Recently, Longstaff, Santa-Clara, and Schwartz (2001) addresses the issue of overfitting in pricing American swaptions. In contrast, under our model, the three factors capture all that is uncertain. We have no other parameters to estimate and hence no other risks to bear. Furthermore, we find that the empirical performance of our model in fitting the term structure of U.S. swap rates and Treasury yields is comparable to the much more complicated models.

The collapse of dimensionality is also observed in the geometric analysis of Pan (1998). In this paper, we link the collapse of dimensionality to the risk-neutral dynamics of the interest rates. To guarantee that all interest rates are bounded below from zero, we start with the assumption that all continuously compounded spot rates are quadratic forms of a finite-dimensional state vector. This setup belongs to the quadratic class of Leippold and Wu (2002). Nevertheless, the resulting term structure behaves as if all spot rates are proportional to one dynamic factor, which follows a special two-parameter square root process. Thus, the final model falls within the affine class of Duffie and Kan (1996) and is very close to the model of Cox, Ingersoll, and Ross (1985). In a way, our model illustrates the inherent link between the affine class and the quadratic class of term structure models.

With some transformation, we can define the three interest rate factors in terms of the level, the maturity, and the curvature of the maximum forward rate. Thus, the model can naturally generate a hump-shaped term structure. Recent evidence supports such a hump shape. For example, Brown and Schaefer (2000) find that, in nearly ten years of daily data on U.S. Treasury STRIPS from 1985 to 1994, the implied two-year forward rate spanning years 24 to 26 is lower than the forward rate for years 14 and 16 on 98.4 percent of occasions. The average difference in these rates is 138 basis points. A similar downward tilt also appears in estimates of forward rates derived from the prices of coupon bonds in the U.S. Treasury market and in the U.K. market for both real and nominal government bonds. Given the initial upward-sloping term structure in most observations, the downward slopes in the very long term imply a hump-shaped term structure for the forward rates.

II. The Model

We fix a filtered complete probability space $\{\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{0\le t\le \mathcal{T}}\}$ that satisfies the usual technical conditions[1] with $\mathcal{T}$ being some finite, fixed time. We assume that the uncertainty of the economy is governed by a finite-dimensional state vector $\mathbf{u}$.

Assumption 1 (Diffusive State Vector) *Under the probability space $\{\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{0\le t\le \mathcal{T}}\}$, the state vector $\mathbf{u}$ is a d-dimensional Markov process in some state space $\mathcal{D} \subset \mathbb{R}^d$, solving the stochastic differential equation:*

$$d\mathbf{u}_t = \mu(\mathbf{u}_t)dt + \Sigma(\mathbf{u}_t)d\mathbf{z}_t, \tag{1}$$

where $\mathbf{z}_t$ is a vector Wiener process in $\mathbb{R}^d$, $\mu(\mathbf{u}_t)$ is an $d\times 1$ vector defining the drift, and $\Sigma(\mathbf{u}_t)$ is an $n\times n$ matrix defining the diffusion of the process. We further assume that $\mu(\mathbf{u}_t)$ and $\Sigma(\mathbf{u}_t)$ satisfy the usual regularity conditions such that the above stochastic differential equation allows a strong solution.

For ease of notation, we assume for now that the process is time homogeneous. For any time $t \in [0, \mathcal{T}]$ and time-of-maturity $T \in [t, \mathcal{T}]$, we assume that the market value at time t of a zero-coupon bond with maturity $\tau = T - t$ is fully characterized by $P(\mathbf{u}_t, \tau)$ and that the instantaneous interest rate, or the short rate, r, is defined by continuity:

$$r_t \equiv \lim_{\tau\downarrow 0} \frac{-\ln P(\mathbf{u}_t, \tau)}{\tau}. \tag{2}$$

[1] For technical details, see, for example, Jacod and Shiryaev (1987).

We further assume that there exists a risk-neutral measure, or a martingale measure, $\mathbb{P}^*$, under which the bond price can be written as

$$P(\mathbf{u}_t, \tau) = \mathbb{E}_t^* \left[\exp\left(-\int_t^T r_s ds \right) \right], \tag{3}$$

where $\mathbb{E}_t^*[\cdot]$ denotes expectation under measure $\mathbb{P}^*$ conditional on the filtration $\mathcal{F}_t$. Under certain regularity conditions, the existence of such a measure is guaranteed by no-arbitrage. The measure is unique when the market is complete.[2]

Let $\mu^*(\mathbf{u}_t)$ denote the drift function of $\mathbf{u}_t$ under measure $\mathbb{P}^*$. The diffusion function $\Sigma(\mathbf{u}_t)$ remains the same under the two measures by virtue of the Girsanov's theorem.

The spot rate of maturity τ is defined as

$$y(\mathbf{u}_t, \tau) \equiv -\frac{1}{\tau} \ln P(\mathbf{u}_t, \tau). \tag{4}$$

The instantaneous forward rate is defined as

$$f(\mathbf{u}_t, \tau) \equiv -\frac{\partial \ln P(\mathbf{u}_t, \tau)}{\partial \tau}. \tag{5}$$

Assumption 2 (Positive Interest Rates) *The spot rates, y, take the following quadratic form of the state vector* $\mathbf{u}$*,*

$$y(\mathbf{u}_t, \tau) = \frac{1}{\tau} \mathbf{u}_t^\top A(\tau) \mathbf{u}_t, \tag{6}$$

where $A(\tau)$ is a positive definite matrix so that all spot rates are bounded from below at zero.

[2] Refer to Duffie (1992) for details.

As the asymmetric part of A has zero contribution to the spot rate, we also assume that A is symmetric with no loss of generality.

In principle, the positivity of interest rates can be guaranteed either through a quadratic form or through an exponential function. However, the exponential family is not consistent with any diffusion dynamics for the state vector (See Björk and Christensen (1999) and Filipović (1999, 2000)). Furthermore, the history of interest rates across the world (witness Switzerland and, in recent times, Japan) shows that we must allow an interest rate of zero to be reachable. Zero is not reachable if interest rates are specified as exponential functions of the state variable, but can be reached under our quadratic specification by letting the state vector $\mathbf{u}$ approach the vector of zeros. The fact that $\mathbf{u}$ can be small argues against the inclusion of linear terms, since the linear term would dominate when the state vector is small, thus potentially allowing negative interest rates.

Proposition 1 (Bond Pricing) *Under the assumptions of diffusion state dynamics in (1) and positive interest rates in (6), the term structure of zero-coupon bonds is given by*

$$P(r_t, \tau) = \exp\left(-c(\tau) r_t\right), \tag{7}$$

where r_t is the instantaneous interest rate and follows a square-root process under the risk-neutral measure $\mathbb{P}^*$,

$$dr_t = -\kappa r_t dt + \sigma \sqrt{r_t} dw_t, \tag{8}$$

with $\kappa \in \mathbb{R}, \sigma \in \mathbb{R}^+$ being constant parameters and w_t being a newly defined scaler Wiener process. The maturity coefficient $c(\tau)$ is determined by the following Riccati equation:

$$c'(\tau) = 1 - \kappa c(\tau) - \frac{1}{2}\sigma^2 c(\tau)^2, \tag{9}$$

with the boundary condition: $c(0) = 0$.

Although we start with a d-dimensional state vector, the dimension of the term structure collapses to one. The proof of the bond pricing formula follows standard argument. We solve for the coefficients $c(\tau)$ by applying the Feynman-Kac formula and the principle of matching.

Proof. Applying the Feynman-Kac formula to the zero price function in (3) yields:

$$r(\mathbf{u})\, P(\mathbf{u},\tau) = \frac{\partial P(\mathbf{u},\tau)}{\partial t} + \mathcal{L}^* P(\mathbf{u},\tau), \tag{10}$$

where $\mathcal{L}^*$ denotes the infinitesimal generator under the risk-neutral measure $\mathbb{P}^*$ and is given by

$$\mathcal{L}^* P(\mathbf{u},\tau) = \left[\frac{\partial P}{\partial \mathbf{u}}\right]^\top \mu^*(\mathbf{u}) + \frac{1}{2} tr\left[\left(\frac{\partial^2 P}{\partial \mathbf{u} \partial \mathbf{u}^\top}\right)\left(\Sigma(\mathbf{u})\Sigma(\mathbf{u})^\top\right)\right].$$

The quadratic specification for the spot rate in (6) implies that the instantaneous interest rate also has a quadratic form:

$$r(\mathbf{u}) = \mathbf{u}^\top A'(0)\, \mathbf{u}. \tag{11}$$

Plugging the quadratic specifications for the spot rate in (6) and for the short rate in (11) into equation (10), we have

$$\begin{aligned} \mathbf{u}^{\top}A'(0)\mathbf{u} &= \mathbf{u}^{\top}A'(\tau)\mathbf{u} - 2\mathbf{u}^{\top}A(\tau)\mu^{*}(\mathbf{u}) \\ &\quad - tr\left[A(\tau)\left(\Sigma(\mathbf{u})\Sigma(\mathbf{u})^{\top}\right)\right] + 2\left[\mathbf{u}^{\top}A(\tau)\Sigma(\mathbf{u})\Sigma(\mathbf{u})^{\top}A(\tau)\mathbf{u}\right], \end{aligned} \tag{12}$$

which should hold for all maturity τ and states $\mathbf{u}$.

To maintain the quadratic nature of the equation in (12), we need the diffusion term $\Sigma(\mathbf{u})$ to be independent of the state vector $\mathbf{u}$. Let $V \equiv \Sigma\Sigma^{\top}$ denote a positive definite symmetric constant matrix. Indeed, via a rotation of indices, we can set $V = I$ with no loss of generality.

Equation (12) becomes

$$\mathbf{u}^{\top}A'(0)\mathbf{u} = \mathbf{u}^{\top}A'(\tau)\mathbf{u} - 2\mathbf{u}^{\top}A(\tau)\mu^{*}(\mathbf{u}) - tr(A(\tau)V) + 2\mathbf{u}^{\top}A(\tau)^{2}V\mathbf{u} \tag{13}$$

Furthermore, to balance the power of the equation, we decompose the drift function μ^{*} into two parts, $\mu^{*}(\mathbf{u}) = \mu_1(\mathbf{u}) - B\mathbf{u}$, where B denotes a constant matrix and is assumed to be symmetric with no loss of generality. The first part $\mu_1(\mathbf{u},t)$ satisfies the equality:

$$-2\mathbf{u}^{\top}A(\tau)\mu_1(\mathbf{u}) - tr(A(\tau)V) = 0. \tag{14}$$

That is, the role of $\mu_1(\mathbf{u})$ is to cancel out the constant term on the right-hand side of equation (13). However, since the drift term $\mu_1(\mathbf{u})$ cannot depend on maturity τ, for the equality (14) to hold, we must be able to factor out the maturity dependence

$$A(\tau) = a(\tau) D, \tag{15}$$

where $a(\tau)$ is a scalar and D is a positive definite symmetric matrix independent of τ. This maturity separation determines the most important result of this article: the collapse of dimensionality.

Given the maturity separation, equation (14) becomes

$$-2\mathbf{u}^\top D\mu_1(\mathbf{u}) - tr(DV) = 0. \tag{16}$$

Equation (13) becomes

$$a'(0)\,\mathbf{u}^\top D\mathbf{u} = a'(\tau)\,\mathbf{u}^\top D\mathbf{u} + a(\tau)\,2\mathbf{u}^\top DB\mathbf{u} + a(\tau)^2\,2\mathbf{u}^\top D^2 V\mathbf{u}. \tag{17}$$

For this equation to hold for all states $\mathbf{u} \in \mathbb{R}^d$, we need

$$a'(0)\,D = a'(\tau)\,D + 2a(\tau)\,DB + 2a(\tau)^2 D^2 V. \tag{18}$$

After rearrangement, we have

$$a'(\tau)\,I = a'(0)\,I - 2a(\tau)\,B - 2a(\tau)^2 DV. \tag{19}$$

Since the equation needs to hold for all elements of the matrix, we must have

$$2B = \kappa I; \quad DV = \frac{1}{4}\nu I. \tag{20}$$

We hence obtain the ordinary differential equation,

$$a'(\tau) = a'(0) - \kappa a(\tau) - \frac{1}{2}\nu a(\tau)^2. \tag{21}$$

Furthermore, let $x = \mathbf{u}^\top D\mathbf{u}$, the zero price can then be written as

$$-\ln P = \mathbf{u}^\top A(\tau)\mathbf{u} = a(\tau)\mathbf{u}^\top D\mathbf{u} = a(\tau)x. \tag{22}$$

Next, given the state vector process

$$d\mathbf{u} = (\mu_1(\mathbf{u}) - B\mathbf{u})\,dt + \sqrt{V}d\mathbf{z},$$

by Itô's lemma, we obtain the process for x under $\mathbb{P}^*$,

$$\begin{aligned} dx &= 2\mathbf{u}^\top D(d\mathbf{u}) + tr(DV)\,dt \\ &= \left(2\mathbf{u}^\top D\mu_1(\mathbf{u}) + tr(DV) - 2\mathbf{u}^\top DB\mathbf{u}\right)dt + 2\mathbf{u}^\top D\sqrt{V}d\mathbf{z} \\ &= -\kappa x dt + \sqrt{\nu x}dw. \end{aligned}$$

We obtain the last equality by applying (16) and by defining a new Wiener process w:

$$dw=\frac{2\mathbf{u}^{\top}D\sqrt{V}d\mathbf{z}}{\sqrt{4\mathbf{u}^{\top}DVD\mathbf{u}}}=\frac{2\mathbf{u}^{\top}D\sqrt{V}d\mathbf{z}}{\sqrt{v\mathbf{u}^{\top}D\mathbf{u}}}=\frac{2\mathbf{u}^{\top}D\sqrt{V}d\mathbf{z}}{\sqrt{vx}}.$$

The instantaneous interest rate is $r_t=a'(0)x_t$. A rescaling of index

$$c(\tau)=a(\tau)/a'(0),\quad \sigma=\sqrt{va'(0)}, \tag{23}$$

gives us

$$-\ln P=c(\tau)r_t, \tag{24}$$

with

$$c'(\tau) = 1-\kappa c(\tau)-\frac{1}{2}\sigma^2 c(\tau)^2 \tag{25}$$

$$dr_t = -\kappa r_t dt+\sigma\sqrt{r_t}dw. \tag{26}$$

The initial condition $c(0)=0$ is determined by the fact that $P(r_t,0)=1$. ■

Under our model, due to the maturity separability, the dimension of the state space collapses to one. Bonds are priced as if there is only one dynamic factor. Furthermore, this one dynamic factor follows a two-parameter square-root process under the risk-neutral measure $\mathbb{P}^*$. We leave the dynamics of this factor under the physical measure $\mathbb{P}$ unspecified. The specification of the physical dynamics can be separately determined to match the time-series properties of interest rates while satisfying the constraints implied by the Girsanov theorem.

The two parameters of the square-root process determine both the risk-neutral dynamics of the single dynamic factor, and the shape of the yield curve via the ordinary differential equation in (25). In our empirical application, we relax the time-homogeneity assumption and allow the two parameters to vary over time so that we can fit the yield curve at each day. Nevertheless, the bonds are priced as if the two parameters are constant. We hence label them as static factors. Therefore, we obtain a three-factor term structure model. However, this three-factor structure is not a result of exogenous specification, but of a collapse of dimensionality due to the seemingly innocuous contention that all rates are bounded below from zero.

Our three-factor model contrasts sharply with traditional three-factor models in that the three factors in our model summarize everything that is uncertain about the shape of the term structure. Traditional three-factor models often contain many parameters in addition to the three factors. The estimates of these parameters often exhibit large standard errors. Therefore, such models are subject to parameter risk. Under our specification, there are no other risk-neutral parameters to be estimated and hence no other risks to be concerned with — except, of course, the risk of the model itself.

Treating κ and σ as constants, we can solve the term structure coefficients $c(\tau)$ analytically:

$$c(\tau) = \frac{2\left(1 - e^{-2\lambda\tau}\right)}{4\lambda - (2\lambda - \kappa)\left(1 - e^{-2\lambda\tau}\right)}, \tag{27}$$

with $\lambda = \frac{1}{2}\sqrt{\kappa^2 + 2\sigma^2}$. We can see immediately that $c(\tau) > 0$ for all $\tau > 0$. Furthermore, since the short rate follows a square-root process, it is bounded below from zero. Therefore, all spot rates are bounded below from zero. Indeed, in our model, all spot rates follow a square-root process.

Although we start with a quadratic specification for the spot rates, the final bond pricing formula says that spot rates are proportional to one dynamic factor. The square-root dynamics of the short rate brings our model very close to the traditional term structure model of Cox, Ingersoll, and Ross (1985). The key difference lies in the absence of a constant term in the drift of the risk-neutral dynamics and the absence of a constant term in the affine structure of the bond yields. A constant term in the affine structure drives the boundary away from zero and hence violates our assumption that all rates are bounded from zero.

We solve the coefficients $c(\tau)$ treating κ and σ as constants. Yet, in our application, we allow the two parameters to vary every day to fit the current yield curve. Thus, there seems to be inconsistency between the two practices. However, the inconsistency is only an illusion since we treat κ and σ not as time-inhomogeneous parameters, but as static factors. We explicitly recognize the risk associated with the time variation of these factors and hedge the risk away by forming portfolios that are first-order neutral to their variation. Due to the low dimensionality of the factor structure, neutrality can be achieved with a maximum of only four instruments. In contrast, in a traditional three-factor model with more than ten parameters, making a portfolio first-order neutral to all parameters and state variables is impractical due to transaction costs.

Our practice is also decisively different from traditional time-inhomogeneous specifications as often applied under the framework of Heath, Jarrow, and Morton (1992). In these specifications, the model parameters are allowed to vary over time in such a way that we can always fit the current observed term structure perfectly. Thus, these models have little to say about the fair pricing of the yield curve. Furthermore, accommodating the whole yield curve often necessitates accepting an infinite dimensional state space, which create difficulties for hedging practices.

III. The Hump-Shaped Forward Rate Curve

The term structure of the long forward rates has been persistently downward sloping (Brown and Schaefer (2000)). Given the initial upward sloping term structure in most observations, the downward slopes in the very long term imply a hump-shaped term structure for the forward rates. Our model captures very nicely the hump shape of the forward rate curve.

We can rotate the system and redefine the three factors explicitly on the hump shape of the forward rate curve. Formally, we let F denote the maximum of the instantaneous forward rate (the peak of the hump), M the maturity at which the forward rate reaches its maximum, and λ some measure of the curvature of the forward rate curve at the maximum. Then, the instantaneous forward rate at maturity τ is given by[3]

$$f(\tau) = F \operatorname{sech}^2 \left[\lambda(\tau - M)\right]. \tag{28}$$

The parameter λ is related to the curvature of the forward rate curve at the maximum by:

$$\delta(M) \equiv \frac{f''(M)}{f(M)} = -2\lambda^2. \tag{29}$$

The new triplet $[F, M, \lambda]$ defines the same term structure as the original triplet $[r, \kappa, \sigma]$. They are linked by,

$$\begin{aligned} F &= r\left(1 + \frac{\kappa^2}{2\sigma^2}\right), \quad M = -\frac{1}{\lambda}\operatorname{arctanh}\left(\frac{\kappa}{2\lambda}\right), \quad \lambda = \frac{1}{2}\sqrt{\kappa^2 + 2\sigma^2}. \\ r &= F \operatorname{sech}^2(\lambda M), \quad \kappa = -2\lambda \tanh(\lambda M), \quad \sigma^2 = 2\lambda^2 \operatorname{sech}^2(\lambda M). \end{aligned} \tag{30}$$

[3]Refer to Appendix A for a derivation.

The new formulation defines the forward rate curve by controlling the exact shape of the curve at the hump. Thus, if we observe a forward rate curve, we can determine the value of the three factors very easily. In our estimation, we model $T \equiv 1/\lambda$ instead of λ, because it has a natural interpretation of time scale.

In contrast, the original triplet of factors $[r, \kappa, \sigma]$ define the risk-neutral dynamics of the short rate. They also define the level, the slope, and the curvature of the forward rate curve at the short end ($\tau = 0$):

$$f'(0) = r, \quad \frac{f'(0)}{f(0)} = -\kappa, \quad \delta(0) = \frac{f''(0)}{f(0)} = \kappa^2 - \sigma^2.$$

Thus, we see clearly how the risk-neutral dynamics of the short rate interacts with the shape of the forward rate curve. The drift parameter κ controls the initial slope of the forward rate curve. The initial curve is upward sloping when κ is negative. On the other hand, the instantaneous volatility term σ contributes to the curvature of the forward rate curve. The larger the variance, the more concave the forward rate curve.

Furthermore, the two points of the forward rate curve at $t = 0$ and $t = M$ are linked by a unit-free quantity $\gamma = \tanh(\lambda M)$:

$$\frac{f(0) - f(M)}{f(M)} = \gamma^2, \quad \frac{\delta(M) - \delta(0)}{\delta(M)} = 3\gamma^2.$$

Based on these observations, the calibration of the forward rate curve is fairly simple. The factors can be directly mapped to the level and shape of the forward rate curve.

Empirical studies (Litterman and Scheinkman (1991) and Heidari and Wu (2003)) have identified three common factors from the U.S. Treasuries and the swap rates. The three common factors represent the level, the slope, and the curvature of the term structure. In our model, we map the level, the slope, and the curvature of the forward rate curve into a consistent dynamic term structure model. We also map them into the risk-neutral dynamics of the underlying dynamic factor.

IV. Fitting the U.S. Treasury Yields and U.S. Dollar Swap Rates

To investigate the model's performance, we calibrate the model to two sets of data. One is U.S. Treasury constant maturity par yields and the other is U.S. dollar swap rates of the same maturities. We investigate the goodness of fit of the model on the two sets of data. We also extract the three factors from the two markets for each day and analyze the time series dynamics of these factors.

A. Data and estimation

We obtain both the swap rate data and the constant maturity Treasury yields from Lehman Brothers. The maturities include two, three, five, seven, ten, 15, and 30 years. The data are weekly (Wednesday) closing mid quotes from December 14th, 1994 to December 28th, 2000 (316 observations).

Table I reports the summary statistics of the swap rates and Treasury par yields. We observe an upward-sloping mean term structure for both swaps and U.S. Treasuries. The standard deviation for both the levels and the first differences exhibit a hump-shaped term structure with the plateau coming at three-year to five-year maturities. Interest rates are highly persistent. The excess skewness and kurtosis estimates are small for both levels and first differences.

We are interested not only in the empirical fit of the model on the yield curves of different markets, but also in the time series properties of the three factors $X \equiv [F, M, T]$ at each date. (The choice of $[F, M, T]$ over $[r, \kappa, \sigma]$ in the estimation is only for numerical stability reasons.) If we can forecast the three factors, we will be able to forecast the yield curve. A natural way to capture both the daily fitting of the cross-section of the term structure and the forecasting of the time series of interest rates is to formulate the framework into a state space system and estimate the system using Kalman (1960) filter.

For the estimation, we assume that the three factors can be forecasted via a simple VAR(1) system:

$$X_t = A + \Phi X_{t-1} + \varepsilon_t, \tag{31}$$

where ε denotes the forecasting residuals. We use this forecasting equation as the state propagation equation, with ε as the state propagation error with covariance matrix Q. We then construct the measurement equations based on the valuation of the par yields on the Treasury and swap market, respectively,

$$S_t(\tau) = h(X_t, \tau) + e_t, \tag{32}$$

where $h(X_t, \tau)$ denotes the model-implied value of the par yield of maturity τ as a function of the factors X_t and e_t denotes the measurement error, which we assume has a covariance matrix of R. Since the U.S. Treasury par bond and the U.S. dollar swap contract both have semi-annual payment intervals, the model-implied par yield is given by

$$h(X_t, \tau) = 200 \frac{1 - P(\tau)}{\sum_{i=1}^{2\tau} P(i/2)}, \tag{33}$$

where $P(\tau)$ denotes the model-implied value of the zero coupon bond (discount factor) and is given in equation (7). Since the measurement equation is nonlinear in the state vectors, we apply the extended Kalman Filter, under which the conditional variance update is based on a first-order Taylor expansion.

The parameters of the state space system include those that control the forecasting time series dynamics and the covariance matrices of the state propagation errors and measurement errors $\Theta \equiv [A, \Phi, Q, R]$. We estimate these parameters using a quasi-likelihood method assuming that the forecasting errors of the par yields are normally distributed. (Please see Appendix B for more details.) In our estimation, we assume that the measurement errors on each series are independent, but bear distinct variance. Thus, R is a diagonal matrix, with each element denoting the goodness of fit on each corresponding series.

Table II reports the estimates (and standard errors in parentheses) of the state space estimation on both the U.S. dollar swap market and the U.S. Treasury market.

B. Model performance

Table III reports the summary properties of the pricing errors on the swaps and Treasury par yields. We define the error as the difference between the market-observed rates and the model-implied rates, in basis points. The fitting is good despite the simple model structure. Overall, the mean absolute error is within a few basis points. The maximum error is only 28 basis points for the swap rates and 41 basis points for the Treasury par yields. An inspection of the error properties across different maturities indicates that the key difficulty of the model lies in fitting interest rates at short maturities (two years). The mean error on the two year rates is -7.5 basis points for swaps and -4.5 for Treasuries, implying that the observed two-year rates are on average lower than those implied by the model.

Figure 1 plots the time series of the pricing errors on the swap rates (left panel) and the Treasury par yields (right panel) at selected maturities: two, five, ten, and 30 years. We observe that except at short maturities, the pricing errors are normally within ten basis points. The magnitude of these pricing errors is comparable to those reported in much more complicated models.

C. Factor dynamics

By applying the state space estimation, we obtain not only the weekly fits on the yield curve, but also the parametric estimates on the dynamics of the three factors. A detailed specification analysis of the factor time series dynamics and the associated analysis of the market price of risk is beyond the scope of this paper. Therefore, we use only a simple VAR(1) specification to summarize the properties of these factors. In what follows, we analyze the time series of the three factors. We compare how the three factors relate to one another and how the two markets differ.

The properties of swap spreads, which we defined as the difference between the swap rate and the constant maturity Treasury par yield, are of great interest to both practitioners and academics. The magnitude of the swap spread reflects the difference in the default risk of the financial sector that quotes LIBOR rates and the U.S. Treasury. In addition, the swap spread may also include a significant liquidity component. The swap markets are a purely contract-driven market, but the interest rates in the Treasury market are often driven by the supply and demand of certain Treasury issuance. In what follows, we analyze the two components in the swap spreads based on our model structure.

C.1. The dynamic level factor

Under our model structure, the level of the yield curve can be represented by the instantaneous short rate r. The left panel of Figure 2 plots the extracted instantaneous interest rate from the swap market (dashed line) and the Treasury market (solid line). The right panel of Figure 2 depicts the difference (swap spread) between the two short rates. The average spread on the two short rates over this sample period is 34.19 basis points. Overall, the two short rates move very closely to each other. However, the swap spread does change over time. Before 1998, the spread is in general within 40 basis points. The spike in the swap spread in late 1998 and early 1999 corresponds to the hedge fund crisis during that time. The swap spread during year 2000 is also unusually high, corresponding to the reduced supply in the U.S. Treasury as a result of the budget surplus at that time. Thus, although the spread spike in early 1999 can be attributed to a credit event, the spread plateau in 2000 is mainly due to liquidity factors.

C.2. The slope and curvature factors

The slope of the forward rate curve is closely related to the drift parameter κ of the short rate risk-neutral dynamics. The slope is positive when κ is negative. In contrast, the instantaneous volatility σ of the short rate dynamics is closely related to the curvature of the forward rate curve. The higher the volatility, the more concave the forward rate curve.

Figure 3 plots the time series of $-\kappa$ (left panel) and σ (right panel) as an illustration of the slope and curvature dynamics of the yield curve. The solid lines depict the factors extracted from U.S. Treasury market and the dashed lines depict the factors from the swap market. The two markets move closely together as their shape (slope and curvature) of the forward rate curves also move together. Furthermore, comparing the time series of the short rate to that of the slope and curvature factors, we see that the slope and curvature factors tend to move in a direction opposite to the level factor. When the short rate is high, the forward rate curve tends to be flat. The two spikes in the slope and curvature time series correspond to the two dips in the short rate.

V. Extensions: Jumps in Interest Rates

Our model is derived under three important assumptions: the positivity of interest rates, a finite state representation, and a diffusion state dynamics. We contend that interest rate positivity is a necessary condition to guarantee no arbitrage, as long as we are allowed to hold cash for free. A finite state representation is also necessary for complete hedging to be feasible in practice in the presence of transaction costs. However, the assumption on pure-diffusion state dynamics is more for convenience and tractability than for reasonability. We do see that interest rates move discontinuously (jumps) every now and then. In this section, we explore whether incorporating a jump component by itself violates the assumptions on positive interest rates and finite state dynamics and if not, how jumps can be incorporated into the state dynamics.

We start with the degenerating case that the jump component has zero weight in the state dynamics. Then, our previous analysis indicates that zero prices can be written as

$$-\ln P(r_t, \tau) = c(\tau) r_t, \tag{34}$$

where the short rate r_t follows a square-root dynamics with a zero mean:

$$dr_t = -\kappa r_t dt + \sigma \sqrt{r_t} dw_t, \tag{35}$$

and the coefficient $c(\tau)$ satisfies a Riccati equation. As we discussed before, this model serves as a special example of a one-factor affine model.

Duffie, Pan, and Singleton (2000) incorporate Poisson jumps in the affine structure. Filipović (2001) incorporates more general jumps in a one-factor affine structure. Since we are dealing with a one factor structure, we consider the more general jump specification in Filipović (2001). Filipović (2001) proves that under the general affine framework, the positive short rate r_t is a CBI-process (Conservative Branching Process with Immigration), uniquely characterized by its its generator

$$\begin{aligned} \mathcal{A}f(x) &= \frac{1}{2}\sigma^2 x f''(x) + \left(a' - \kappa x\right) f'(x) \\ &\quad + \int_{\mathbb{R}_+^0} \left(f(x+y) - f(x) - f'(x)(1 \wedge y)\right)(m(dy) + x\mu(dy)), \end{aligned} \tag{36}$$

where $a' = a + \int_{\mathbb{R}_+^0} (1 \wedge y)\, m(dy)$ for some numbers $\sigma^2, a \in \mathbb{R}_+, \kappa \in \mathbb{R}$ and non-negative Borel measures $m(dy)$ and $\mu(dy)$ on $\mathbb{R}_+^0$ (the positive real line excluding zero) satisfying

$$\int_{\mathbb{R}_+^0} (1 \wedge y)\, m(dy) + \int_{\mathbb{R}_+^0} \left(1 \wedge y^2\right) \mu(dy) < \infty. \tag{37}$$

We can obtain our current model by setting the jump part to zero and the constant part of the drift of the square root process to zero ($a = 0$). The two Borel measures define two jump components. The jump component defined by $m(dy)$ is a direct addition to the diffusion process. The jump component defined by $\mu(dy)$ is specified as proportional to x. Hence, we label the former as a constant jump component and the latter a proportional jump component. In essence, the arrival rate of jumps in the "constant" component does not depend on the short rate level, but the arrival rate of the "proportional" component is proportional to the short rate level. Condition (37) requires that the jump component defined by $m(dy)$ exhibit finite variation and the jump component defined by $\mu(dy)$ exhibit finite quadratic variation.

Under the specification in (36), the zero prices are given by

$$-\ln P(r_t, \tau) = A(\tau) + B(\tau) r_t \tag{38}$$

with $A(\tau)$ and $B(\tau)$ solve uniquely the generalized Riccati equations

$$\begin{aligned} B'(\tau) &= R(B(\tau)), \quad B(0) = 0 &(39)\\ A(\tau) &= \int_0^\tau F(B(s))\,ds, &(40) \end{aligned}$$

where R and F are defined as

$$\begin{aligned} R(\lambda) &\equiv 1 - \kappa\lambda - \frac{1}{2}\sigma^2\lambda^2 + \int_{\mathbb{R}_+^0} \left(1 - e^{-\lambda y} - \lambda(1 \wedge y)\right)\mu(dy); &(41)\\ F(\lambda) &\equiv a\lambda + \int_{\mathbb{R}_+^0} \left(1 - e^{-\lambda y}\right) m(dy). \end{aligned}$$

To guarantee that all rates are bounded from zero, we need to set $A(\tau) = 0$ for all τ, which we obtain by setting $a = 0$ and $m(dy) = 0$. The condition $a = 0$ is already known. The second condition $m(dy) = 0$ says that we cannot add a constant jump component while maintaining that all rates are bounded from zero. Nevertheless, we can incorporate a proportional jump component. Since $B(\tau)$ is positive for all τ, all interest rates are bounded from zero. In absence of the proportional jump component, $R(\lambda)$ is reduced to our Riccati equation for the diffusion case. The last term in (41) captures the contribution of the proportional jump component.

VI. Conclusion

In this paper, we contend that all interest rates should be bounded from below at zero. Such a seemingly innocuous contention, together with the assumption of continuity, results in a dramatic collapse of dimensionality. Such conditions lead to a term structure model that has only one dynamic factor and two static factors. Even more surprising, there are no other parameters in the model that affect the shape of the term structure. Therefore, model calibration becomes a trivial problem and there no longer exists a distinction between out-of-sample and in-sample performance. Furthermore, risks from the three factors can be hedged away easily with only a few instruments. Since there are no more parameters, the model is not subject to any parameter risk.

To put the model into practical application, we cast the model in a state space framework and estimate the three states via quasi maximum likelihood together with an extended Kalman filter. We apply this estimation procedure to both the U.S. Treasury market and the U.S. dollar swap market. Despite its extreme simplicity, the model performs well in fitting the daily term structures of both markets. A time series analysis of the extracted factors from the two markets provides us with some interesting insights on the evolution of the interest rate market.

A potential application of the model, which can be explored in future research, is to forecast the term structure of interest rates. Recently, Diebold and Li (2001) illustrate how the Nelson-Siegel framework can be applied successfully to forecasting the term structure of Treasury yields. Yet, the inherent inconsistency of the Nelson-Siegel model is well-documented in Björk and Christensen (1999) and Filipović (1999, 2000). Our model provides a parsimonious but consistent alternative to the Nelson-Siegel framework.

Appendix A. Factor Representation

The term structure is determined by the following ordinary differential equation:

$$c'(\tau) = 1 - \kappa c(\tau) - \frac{1}{2}\sigma^2 c(\tau)^2, \tag{A1}$$

with $c(0) = 0$. One solution of this Riccati equation is given in (27). Another way of solving the equation is through the following change of variables:

$$\psi(\tau) \equiv \frac{c(\tau)\sigma^2 + \kappa}{\sqrt{\kappa^2 + 2\sigma^2}}, \quad \lambda = \frac{1}{2}\sqrt{\kappa^2 + 2\sigma^2}, \tag{A2}$$

where $\kappa^2 + 2\sigma^2$ defines the discriminant of the ordinary differential equation. Then the ordinary differential equation (A1) is transformed into the elementary problem

$$\psi'(\tau) = \lambda\left(1 - \psi(\tau)^2\right) \tag{A3}$$

with $\psi(0) = \kappa/(2\lambda)$.

The solution of (A3) is

$$\psi(\tau) = \tanh\left[\lambda\left(\tau - M\right)\right],$$

where M is defined by the boundary condition

$$\psi(0) = \tanh\left(-\lambda M\right) = \frac{\kappa}{2\lambda}.$$

That is,

$$M = -\frac{1}{\lambda}\text{arctanh}\left(\frac{\kappa}{2\lambda}\right).$$

Translating $\psi(\tau)$ back to the bond pricing coefficients $c(\tau)$ gives

$$c(\tau) = \frac{2\lambda}{\sigma^2}\left[\tanh\lambda(t-M) + \tanh\lambda M\right] \tag{A4}$$

The instantaneous forward rate is given by

$$f(\tau) = c'(\tau)r = \frac{2\lambda^2 r}{\sigma^2}\text{sech}^2\lambda(\tau - M) = F\,\text{sech}^2\lambda(\tau - M)$$

where

$$F = \frac{2\lambda^2 r}{\sigma^2} = r\left(1 + \frac{\kappa^2}{2\sigma^2}\right),$$

is the maximal forward rate and M is the corresponding maturity.

Appendix B. Extended Kalman Filter and Quasi Likelihood

The state space estimation method is based on a pair of state propagation and measurement equations. In our application, the state vector X propagates according to VAR(1) processes specified in (31). The measurement equation is given in (32), which is based on the valuation of the par yield. Let $\overline{X}_t$ denote the *a priori* forecast of the state vector at time t conditional on time $t-1$ information and $\overline{V}_t$ the corresponding conditional covariance matrix. Let $\widehat{X}_t$ denote the *a posteriori* update on the time t state vector based on observations (S_t) at time t and $\widehat{V}_t$ the corresponding *a posteriori* covariance matrix. Then, based our OU state process specification, the state propagation equation is linear and Gaussian. The *a priori* update equations are:

$$\begin{aligned} \overline{X}_t &= A + \Phi\widehat{X}_{t-1}; \\ \overline{V}_t &= \Phi\widehat{V}_{t-1}\Phi^\top + Q. \end{aligned} \tag{B5}$$

The filtering problem then consists of establishing the conditional density of the state vector X_t, conditional on the observations up to and including time t. In case of a linear measurement equation,

$$S_t = HX_t + e_t,$$

the Kalman Filter provides the efficient *a posteriori* update on the conditional mean and variance of the state vector:

$$\begin{aligned}
\overline{S}_t &= H\overline{X}_t; \\
\overline{A}_t &= H\overline{V}_t H^\top + R \\
K_t &= \overline{V}_t H \left(\overline{A}_t\right)^{-1}; \\
\widehat{X}_t &= \overline{X}_t + K_t \left(S_t - \overline{S}_t\right); \\
\widehat{P}_t &= (I - K_t H)\overline{V}_t,
\end{aligned} \tag{B6}$$

where $\overline{S}_t$ and $\overline{A}_t$ are the *a priori* forecasts on the conditional mean and variance of the observed series and R are the covariance matrix of the measurement errors.

However, in our application, the measurement equation in (32) is nonlinear. We apply the Extended Kalman Filter (EKF), which approximates the nonlinear measurement equation with a linear expansion:

$$S_t \approx H\left(\overline{X}_t\right) X_t + e_t, \tag{B7}$$

where

$$H\left(\overline{X}_t\right) = \left. \frac{\partial h\left(\overline{X}_t\right)}{\partial X_t} \right|_{X_t = \overline{X}_t}. \tag{B8}$$

Thus, although we still use the original pricing relation to update the conditional mean, we update the conditional variance based on this linearization. For this purpose, we need to numerically evaluate the derivative defined in (B8). We follow Norgaard, Poulsen, and Raven (2000) in updating the Cholesky factors of the covariance matrices directly.

Using the state and measurement updates, we obtain the one-period ahead forecasting error on the par yields,

$$e_t = S_t - \overline{S}_t = S_t - h\left(\overline{X}_t\right).$$

Assuming that the forecasting error is normally distributed, the quasi log-likelihood function is given by

$$\mathcal{L}(\mathbf{S}) = \sum_{t=1}^{T} l_t, \tag{B9}$$

where

$$l_t = -\frac{1}{2}\log\left|\overline{A}_t\right| - \frac{1}{2}\left(e_t^{\top}\left(\overline{A}_t\right)^{-1} e_t\right),$$

where the conditional mean $\overline{S}_t$ and variance $\overline{A}_t$ are given in the EFK updates in (B6).

REFERENCES

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, Quadratic term structure models: Theory and evidence, *Review of Financial Studies* 15, 243–288.

Babbs, S. H., 1997, Rational bounds, Working paper, First National Bank of Chicago.

Backus, David, Silverio Foresi, Abon Mozumdar, and Liuren Wu, 2001, Predictable changes in yields and forward rates, *Journal of Financial Economics* 59, 281–311.

Balduzzi, Pierluigi, Sanjiv Das, Silverio Foresi, and Rangarajan Sundaram, 1996, A simple approach to three-factor affine term structure models, *Journal of Fixed Income* 6, 43–53.

Beaglehole, D. R., and M.S. Tenney, 1991, General solution of some interest rate-contingent claim pricing equations, *Journal of Fixed Income* 1, 69–83.

Beaglehole, D. R., and M.S. Tenney, 1992, A nonlinear equilibrium model of term structures of interest rates: Corrections and additions, *Journal of Financial Economics* 32, 345–454.

Björk, Tomas, and Bent Jesper Christensen, 1999, Interest rate dynamics and consistent forward rate curves, *Mathematical Finance* 9, 323–348.

Black, Fischer, 1995, Interest rates as options, *Journal of Finance* 50, 1371–1376.

Brandt, Michael, and David A. Chapman, 2002, Comparing multifactor models of the term structure, Working paper, Duke University.

Brandt, Michael, and Amir Yaron, 2001, Time-consistent no-arbitrage models of the term structure, Working paper, University of Pennsylvania.

Brown, Roger H., and Stephen M. Schaefer, 2000, Why long term forward interest rates (almost) always slope downwards, Working paper, Warburg Dillion Read and London Business School UK.

Chen, Ren-Raw, and Louis Scott, 1993, Maximum likelihood estimation of a multifactor equilibrium model of the term structure of interest rates, *Journal of Fixed Income* 3, 14–31.

Constantinides, G. M., 1992, A theory of the nominal term structure of interest rates, *Review of Financial Studies* 5, 531–552.

Cox, John C., Jonathan E. Ingersoll, and Stephen R. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

Dai, Qiang, and Kenneth Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Dai, Qiang, and Kenneth Singleton, 2002, Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415–441.

Dai, Qiang, and Kenneth Singleton, 2003, Term structure dynamics in theory and reality, *Review of Financial Studies* 16, 631–678.

Diebold, Francis X., and Canlin Li, 2001, Modeling and forecasting the term structure of government bond yields, Working paper, University of Pennsylvania.

Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405–443.

Duffie, Darrell, 1992, *Dynamic Asset Pricing Theory*. (Princeton University Press Princeton, New Jersey) second edn.

Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379–406.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump diffusions, *Econometrica* 68, 1343–1376.

El Karoui, Nicole, R. Myneni, and R. Viswanathan, 1992, Arbitrage pricing and hedging of interest rate claims with state variables: I theory, working paper, University of Paris.

Filipović, Damir, 1999, A note on the Nelson-Siegel family, *Mathematical Finance* 9, 349–359.

Filipović, Damir, 2000, Exponential-polynomial families and the term structure of interest rates, *Bernoulli* 6, 1–27.

Filipović, Damir, 2001, A general characterization of one factor affine term structure models, *Finance and Stochastics* 5, 389–412.

Flesaker, B., and L.P. Hughston, 1996, Positive interest, *RISK* 9, 46–49.

Goldstein, Robert, and William Keirstead, 1997, On the term structure of interest rates in the presence of reflecting and absorbing boundaries, Working paper, Ohio-State University.

Gorovoi, Viatcheslav, and Vadim Linetsky, 2003, Black's model of interest rates as options, eigenfunction expansions and japanese interest rates, *Mathematical Finance* forthcoming.

Heath, David, Robert Jarrow, and Andrew Morton, 1992, Bond pricing and the term structure of interest rates: A new technology for contingent claims valuation, *Econometrica* 60, 77–105.

Heidari, Massoud, and Liuren Wu, 2003, Are interest rate derivatives spanned by the term structure of interest rates?, *Journal of Fixed Income* 13, 75–86.

Jacod, Jean, and Albert N. Shiryaev, 1987, *Limit Theorems for Stochastic Processes*. (Springer-Verlag Berlin).

Jamshidian, F., 1996, Bond, futures and option valuation in the quadratic interest rate model, *Applied Mathematical Finance* 3, 93–115.

Jin, Yan, and Paul Glasserman, 2001, Equilibrium positive interest rates: A unified view, *Review of Financial Studies* 14, 187–214.

Kalman, Rudolph Emil, 1960, A new approach to linear filtering and prediction problems, *Transactions of the ASME–Journal of Basic Engineering* 82, 35–45.

Knez, Peter J., Robert Litterman, and Jose Scheinkman, 1994, Explorations into factors explaining money market returns, *Journal of Finance* 49, 1861–1882.

Leippold, Markus, and Liuren Wu, 2002, Asset pricing under the quadratic class, *Journal of Financial and Quantitative Analysis* 37, 271–295.

Leippold, Markus, and Liuren Wu, 2003, Design and estimation of quadratic term structure models, *European Finance Review* 7, 47–73.

Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 1, 54–61.

Longstaff, Francis A., 1989, A nonlinear general equilibrium model of the term structure of interest rates, *Journal of Financial Economics* 23, 195–224.

Longstaff, Francis A., Pedro Santa-Clara, and Eduardo S. Schwartz, 2001, Throwing away a million dollars: The cost of suboptimal exercise strategies in the swaptions market, *Journal of Financial Economics* forthcoming.

Norgaard, Magnus, Niels K. Poulsen, and Ole Raven, 2000, New developments in state estimation for nonlinear systems, *Automatica* 36, 1627–1638.

Pan, Enlin, 1998, Collpase of detail, *International Journal of Theoretical and Applied Finance* 1, 247–282.

Rogers, L. C. G., 1995, *Mathematical Finance* vol. IMA Volume 65. (Springer New York).

Rogers, L. C. G., 1996, Gaussian errors, *RISK* 9, 42–45.

Rogers, L. C. G., 1997, The potential approach to the term structure of interest rates and foreign exchange rates, *Mathematical Finance* 7, 157–176.

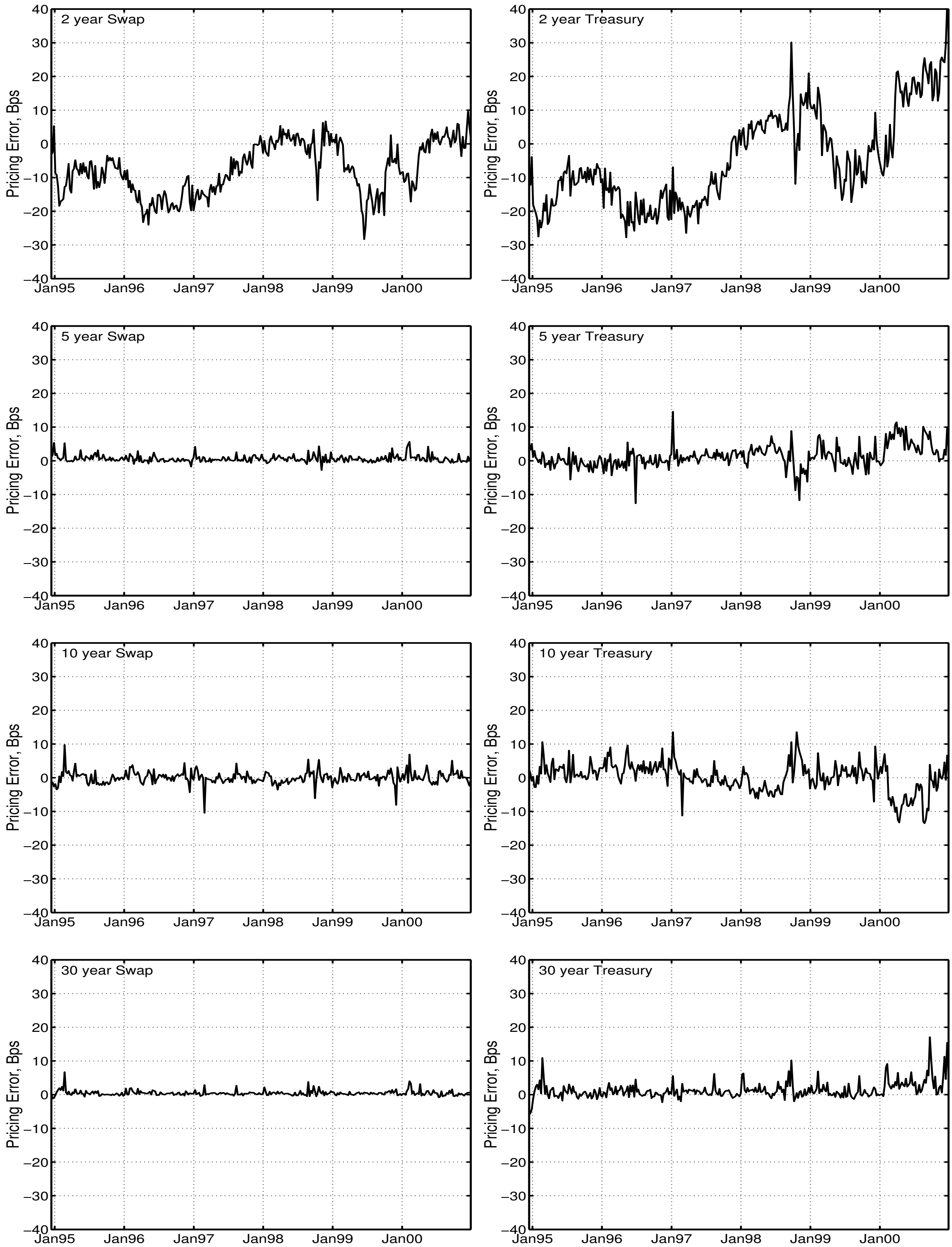

Figure 1. Swap Rate Pricing Errors

Lines report the time series of the pricing errors on swap rates and Treasury par yields. The pricing error is in basis points, defined as the difference between the market-observed rate and the model-implied rate. We compute the errors based on the model estimates in Table II.

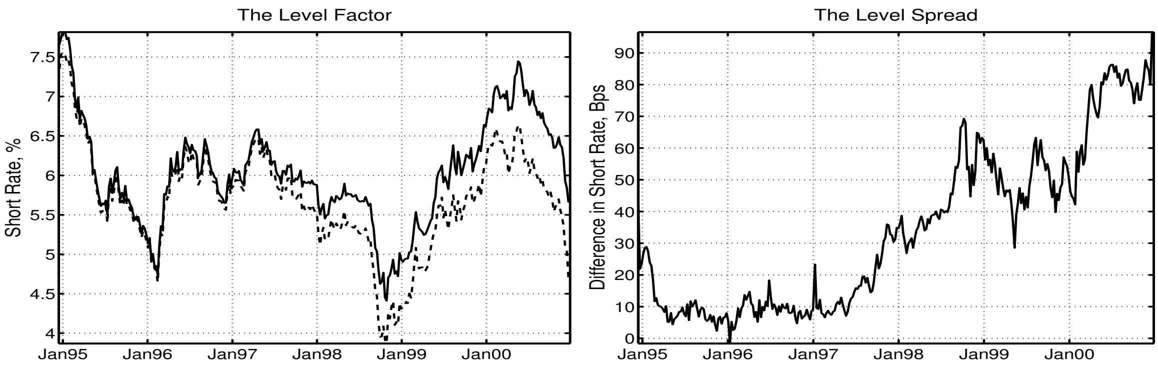

Figure 2. The Short Rate and Swap Spreads

The left panel depicts the instantaneous interest rate (in percentages) implied from the U.S. Treasury market (solid line) and the swap market (dashed line). The right panel depicts the spread, in basis points, between the short rate from the swap market and the short rate from the Treasury market.

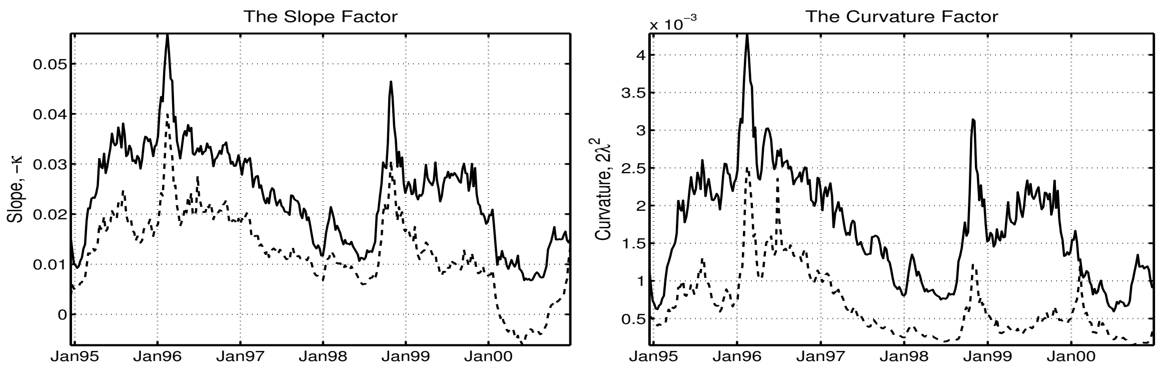

Figure 3. The Slope Factor $-\kappa$ and the Curvature Factor σ

Lines depict the slope factor ($-\kappa$, left panel) and the curvature factor (σ, right panel) extracted from the U.S. Treasury market (solid line) and the swap market (dashed line).

Table I
Summary Statistics of U.S. Dollar Swap Rates and U.S. Treasury Par Yields

	Swap					Treasury				
Mat	Mean	Std	Skew	Kurt	Auto	Mean	Std	Skew	Kurt	Auto
					Levels					
2	6.190	0.656	0.343	0.293	0.971	5.799	0.631	-0.050	0.858	0.969
3	6.303	0.657	0.266	0.224	0.971	5.857	0.654	-0.071	0.798	0.969
5	6.454	0.642	0.133	0.004	0.971	5.947	0.681	-0.143	0.672	0.969
7	6.560	0.629	0.061	-0.147	0.971	5.994	0.672	-0.081	0.522	0.970
10	6.681	0.615	-0.022	-0.300	0.971	6.059	0.669	-0.008	0.238	0.971
15	6.817	0.591	-0.056	-0.388	0.969	6.111	0.653	0.052	0.152	0.971
30	6.889	0.576	0.050	-0.247	0.971	6.266	0.624	0.243	-0.207	0.974
					Differences					
2	-0.007	0.120	0.328	0.977	0.027	-0.008	0.119	0.201	0.642	0.001
3	-0.007	0.123	0.349	0.834	0.020	-0.008	0.121	0.316	0.746	-0.013
5	-0.007	0.123	0.214	0.627	0.013	-0.009	0.124	0.178	0.600	0.008
7	-0.007	0.122	0.264	0.593	0.019	-0.009	0.120	0.282	0.592	0.013
10	-0.007	0.119	0.188	0.410	0.043	-0.009	0.118	0.233	0.353	0.040
15	-0.007	0.117	0.304	0.393	-0.024	-0.008	0.114	0.348	0.319	-0.015
30	-0.007	0.106	0.445	0.545	0.018	-0.008	0.102	0.415	0.528	0.033

The table presents summary statistics of U.S. dollar swap rates and U.S. Treasury par yields. Mean, Std, Skew, Kurt, and Auto denote, respectively, the sample estimates of the mean, standard deviation, skewness, kurtosis, and first-order autocorrelation. The data are weekly closing mid quotes from Lehman Brothers, from December 14th, 1994 to December 28th, 2000 (316 observations).

Table II
Summary Statistics of the Three Factors from Swaps and US STRIPS

Data	Swap	Treasury
State Propagation Equation: $X_t = A + \Phi X_{t-1} + \varepsilon_t, \quad S_\varepsilon S_\varepsilon^\top = Cov(\varepsilon)$		
$A =$	$\begin{bmatrix} 0.1831 & (0.1233) \\ 3.0572 & (0.6830) \\ 7.9858 & (1.4326) \end{bmatrix}$	$\begin{bmatrix} 0.2405 & (1.4971) \\ 4.0164 & (2.2390) \\ 4.4539 & (3.9432) \end{bmatrix}$
$\Phi =$	$\begin{bmatrix} 0.9761 & -0.0052 & -0.0027 \\ (0.0133) & (0.0307) & (0.0081) \\ -0.0163 & 0.9005 & -0.0110 \\ (0.0081) & (0.0164) & (0.0046) \\ -0.0386 & -0.1895 & 0.9350 \\ (0.0164) & (0.0425) & (0.0115) \end{bmatrix}$	$\begin{bmatrix} 0.9660 & 0.0018 & -0.0046 \\ (0.1497) & (0.0057) & (0.0051) \\ -0.0358 & 0.9673 & -0.0122 \\ (0.0303) & (0.0123) & (0.0106) \\ 0.0136 & -0.0286 & 0.9299 \\ (0.0490) & (0.0233) & (0.0202) \end{bmatrix}$
$S_\varepsilon =$	$\begin{bmatrix} 1.1173 & 0 & 0 \\ (0.0552) & & \\ -0.0634 & 0.7012 & 0 \\ (0.0577) & (0.0413) & \\ -0.7486 & 0.3839 & 1.5043 \\ (0.1723) & (0.1862) & (0.1305) \end{bmatrix}$	$\begin{bmatrix} 1.1418 & 0 & 0 \\ (0.0726) & & \\ 0.1898 & 2.3490 & 0 \\ (0.2869) & (0.2455) & \\ -1.5145 & 2.3074 & 3.6332 \\ (0.6367) & (0.6562) & (0.3997) \end{bmatrix}$
Measurement Equation: $S_t = h(X_t) + e_t, \quad Cov(e) = diag\left[\sigma_i^2\right], \quad i = 2,3,5,7,10,15,30.$		
$\begin{bmatrix} \sigma_2 \\ \sigma_3 \\ \sigma_5 \\ \sigma_7 \\ \sigma_{10} \\ \sigma_{15} \\ \sigma_{30} \end{bmatrix} =$	$\begin{bmatrix} 0.1106 & (0.0115) \\ 0.0512 & (0.0051) \\ 0.0114 & (0.0012) \\ 0.0130 & (0.0007) \\ 0.0188 & (0.0007) \\ 0.0295 & (0.0017) \\ 0.0127 & (0.0026) \end{bmatrix}$	$\begin{bmatrix} 0.1468 & (0.0180) \\ 0.0928 & (0.0112) \\ 0.0314 & (0.0016) \\ 0.0003 & (0.0785) \\ 0.0409 & (0.0021) \\ 0.0498 & (0.0039) \\ 0.0285 & (0.0027) \end{bmatrix}$
$\mathcal{L}(\times 10^{-3})$	5.6517	4.7706

The table reports the parameter estimates (standard deviations in parentheses) of the state space system. The state propagation captures the dynamics of the three factors $X_t \equiv [F_t, M_t, T_t]$, where F_t is represented in one thousandth, and M and T are in years. The standard deviation of the measurement error (σ_i) captures the model's performance in fitting the constant maturity yields or swap rates of the denoted maturities. The standard deviation is measured in annual percentages. The model is calibrated to both the U.S. dollar swap rates (left panel) and the U.S. Treasury constant maturity par yields, both of which are weekly data from December 14th, 1994 to December 28th, 2000 (316 observations).

Table III
Summary Statistics of Pricing Errors on U.S. Dollar Swap rates and U.S. Treasury Par Yields

	Swap					Treasury				
Mat	Mean	Std	Mae	Max	Auto	Mean	Std	Mae	Max	Auto
2	-7.524	7.641	8.611	28.290	0.893	-4.358	13.970	12.669	41.425	0.923
3	-2.681	4.053	3.948	13.102	0.751	-1.731	9.243	8.077	31.094	0.871
5	0.608	1.053	0.796	5.635	0.158	1.327	3.400	2.602	14.568	0.531
7	0.843	1.323	1.087	7.859	0.257	0.723	1.401	0.857	8.929	0.111
10	0.022	1.837	1.279	10.430	0.245	0.249	4.315	3.221	13.536	0.674
15	-0.879	2.446	2.052	8.118	0.629	-3.423	2.983	3.947	12.308	0.434
30	0.445	0.763	0.554	6.676	0.160	1.341	2.503	1.758	17.064	0.468

The table presents summary statistics of the pricing errors on U.S. dollar swap rates and U.S. Treasury par yields. We define the pricing error as the difference, in basis points, between the market observed rates and the model implied rates. Mean, Std, Mae, Max, and Auto denote, respectively, the sample estimates of the mean, standard deviation, mean absolute error, max absolute error, and first-order autocorrelation. The market observed rates are weekly closing mid quotes from Lehman Brothers, from December 14th, 1994 to December 28th, 2000 (316 observations). We compute the model-implied rates based on the state space system estimated in Table II.