11-1999

Design and Estimation of Affine Yield Models

David Backus
New York University

Chris I. Telmer
Carnegie Mellon University, chris.telmer@cmu.edu

Liuren Wu
Fordham University

Follow this and additional works at: http://repository.cmu.edu/tepper
Part of the Economic Policy Commons, and the Industrial Organization Commons
Design and Estimation of Affine Yield Models*

David Backus,† Chris Telmer,‡ and Liuren Wu§

First draft: June 4, 1999
This version: November 28, 1999

Abstract

We consider the design and estimation of affine term structure models, starting with a list of properties of bond yields one might want a model to reproduce. We emphasize one property that we think is particularly informative about model structure: the hump-shaped dynamics of bond yields of most maturities. We estimate a model that reproduces this property and some others, and show how it might account for the typical hump-shaped volatility term structure for interest rate caps.

JEL Classification Codes: E43, G12.

Keywords: pricing kernels; stochastic volatility; leptokurtosis; interest rate caps and floors.

---

*We thank Frank Diebold, Rohit Deo and Stan Zin for valuable comments and advice. The most recent version of this paper is available at http://bertha.gsia.cmu.edu.
†Stern School of Business, New York University and NBER; dbackus@stern.nyu.edu.
‡Graduate School of Industrial Administration, Carnegie Mellon University; telmerc@andrew.cmu.edu.
§Graduate School of Business, Fordham University; liwu@mary.fordham.edu.
Design and Estimation of Affine Yield Models

Abstract

We consider the design and estimation of affine term structure models, starting with a list of properties of bond yields one might want a model to reproduce. We emphasize one property that we think is particularly informative about model structure: the hump-shaped dynamics of bond yields of most maturities. We estimate a model that reproduces this property and some others, and show how it might account for the typical hump-shaped volatility term structure for interest rate caps.
1 Introduction and Summary

One of the appealing features of the modern theory of bond pricing is the unifying role played by the pricing kernel: the stochastic process for the kernel governs the behavior of bond yields and prices of fixed income securities in general. When we look at data, however, this unity places enormous demands on the ability of models to account for the complex interactions among bond yields and prices of related derivatives. We start with properties of interest rate data and ask: What kind of model is needed to account for the observed dynamics of the yield curve?

We focus on the popular class of affine models, in which bond yields are linear functions of a vector of state variables. Examples with this form have been studied for decades, but Duffie and Kan’s (1996) characterization of the complete class has served as a catalyst for more systematic study of model design. The leaders in this effort have been Dai and Singleton (1997), who show how generalizations of earlier models lead to substantial improvements in their ability to account for dollar swap rates.

In designing a model, we employ what Leamer (1978, p 11) calls “Sherlock Holmes inference,” in which prominent features of the data guide the model design process. The name stems from Holmes’ inclination to collect evidence before constructing a theory of a crime. Leamer quotes him as telling Watson: “No data yet. ... It is a capital mistake to theorize before you have all the evidence.” This turns the textbook approach to inference on its head, but given the amount of work already done on the limited sample of interest rate data, we think it’s inevitable. It also highlights the impact of different pieces of evidence on model structure.

The most significant properties of US government bond yields since 1970, in our view, include the average slope and shape of the yield curve, persistence in yields of all maturities, persistence in conditional variances of yields, leptokurtosis in yield changes, and hump-shaped dynamics in yields of most maturities. We document each of these properties, imposing as little a priori structure as possible. In this respect we have followed Holmes’ advice to collect evidence first.

With facts in hand, we turn to models. The one-factor Cox-Ingersoll-Ross model serves as a convenient benchmark: its weaknesses highlight properties of the data that a better model might address. In the Cox-Ingersoll-Ross model, the mean yield curve exhibits too little concavity, yields are too highly (in fact, perfectly) correlated across maturities, yield changes exhibit too little kurtosis, and interest rate dynamics
are monotonic, not hump-shaped. Multi-factor models perform better along some dimensions, but the details are critical. We argue that a reasonable model must incorporate dynamic interaction among factors and allow interest rate innovations to be nonnormal. An example illustrates how this might be done. An application to interest rate caps shows how the typical hump in the term structure of volatility might be attributed to a combination of hump-shaped dynamics in bond yields and (perhaps more surprising) leptokurtosis in the conditional distribution of interest rate changes.

2 Evidence

We begin by summarizing some of the salient features of US government bond yields over the last 25 years. The data are monthly, 1970 to 1995, computed from US treasury prices by the “smoothed Fama-Bliss” method using programs supplied by Robert Bliss. In what follows, \( b_t^n \) is the dollar price at date \( t \) of a claim to one dollar at \( t + n \), with both \( t \) and \( n \) measured in months. The continuously-compounded yield or spot rate on an \( n \)-period discount bond at date \( t \) by \( y_t^n \) is defined by

\[
y_t^n = -n^{-1} \log b_t^n.
\]

The short rate is \( r_t = y_t^1 \).

The fixed income literature is filled with studies of interest rate data. We focus on five “facts” that we think have been robust and continuing features of US interest rates and that bear directly on the structure of interest rate models.

**Fact 1.** Average bond yields are an increasing, concave function of maturity. To be sure, bond yields often decline with maturity between 10 and 30 years, but the average yield curve invariably increases sharply at short maturities and rises more slowly between 2 and 10 years. Table 1 is an example: the mean yield in our sample increases 114 basis points between 1 month and 2 years but only 71 basis points over the next 8 years.

**Fact 2.** Yields and yield spreads are highly persistent. Autocorrelations of yields and yield spreads indicate substantial persistence (Tables 1 and 2). Yields, in particular, have first-order autocorrelations above 0.95, leading some observers to suggest a unit root. Equally interesting, in our view, is that the autocorrelations differ: long rates are more persistent than short rates, which in turn are substantially
more persistent than spreads. Figure 1 makes this point for the autocorrelation functions of the 1-month yield, the 5-year yield, and the 2-year spread over the short rate.

**Fact 3. Conditional variances of yields are persistent.** Changes in yields, like changes in most asset prices, exhibit substantial persistence in their conditional variances. Although conditional variances are not directly observable, persistence is evident in the autocorrelations of squared residuals of bond yields from third-order autoregressions; see Figure 2. The autoregressions are an attempt to eliminate predictable variation in conditional means. In our experience, how this is done has little impact on the result.

With more structure, we get a clearer picture of the behavior of conditional variances. Consider the “RiskMetrics” conditional variance estimator,

\[ h_t = \varphi h_{t-1} + (1 - \varphi) \eta_t^2, \]

where \( \eta \) is the residual from the third-order autoregression. We use \( \varphi = 0.95 \) and set \( h_0 \) equal to the sample variance, but little of what follows depends on these choices. Figure 3 suggests that conditional standard deviations of different maturities have largely moved up and down together. The correlation is 0.79 for the 1-month and 10-year rates and higher for other pairs of maturities. Figure 4 suggests that the conditional variance of the short rate is correlated with the short rate, but the correlation is far from perfect (about 0.6). Other maturities (not reported) exhibit similar behavior. In our view, the figure casts doubt on one-factor interest-rate models like those studied by Ait-Sahalia (1996) and Chan, Karolyi, Longstaff, and Sanders (1992), in which the conditional variance is a deterministic function of the short rate. One might be tempted to attribute the imperfect correlation to the inevitable noise in the estimate of \( h \), but more formal studies by Andersen and Lund (1997) and Brenner, Harjes, and Kroner (1996) come to the same conclusion. Even more surprising, the correlation between \( h \) and \( r \) is just as high (0.6) for the post-1985 period as it is for the full sample, which includes the early-80s period of exceptionally high values for both.

**Fact 4. Yield changes and “residuals” are leptokurtotic.** Changes in yields and residuals from autoregressions exhibit substantial leptokurtosis at most maturities (Tables 2 and 3). In this respect, interest rates resemble equity and currency prices. Equally important for model design: at least some of the kurtosis seems to be left once we account for the random behavior of volatility. To see this, suppose we divide residuals into the product of the conditional standard deviation
and a standardized residual $\varepsilon$:

$$\eta_t = h_{t-1}^{1/2}\varepsilon_t.$$  

Even if $\varepsilon$ were normal, randomness in $h$ would impart excess kurtosis to $\eta$. In fact, when we estimate $\varepsilon$ using our RiskMetrics estimate of $h$, we find substantial kurtosis remaining (Table 3).

**Fact 5. Yield dynamics are hump-shaped.** Although it is rarely mentioned, we find that yields exhibit hump-shaped dynamics, a feature of the data that shows up in a variety of ways. A relatively simple one is based on univariate estimates of ARMA models, which we regard as convenient summaries of interest rate dynamics. For each maturity, we estimate ARMA($p,q$) models for orders $0 \leq p,q \leq 3$. We then choose the “best” such models based on the Schwartz and Akaike criteria, respectively. (See Diebold 1998, pp 85-91.) The orders of the best models are reported in Table 4. The Schwartz criterion imposes a larger penalty for extra parameters, and thus leads to more parsimonious models, but the qualitative properties of the two sets of models are similar. Both imply hump-shaped dynamics: for all maturities but the short rate, the impulse response functions are initially increasing (see Table 4 and Figure 4).

Another perspective on yield dynamics comes from multiperiod differences: changes $y_t - y_{t-k}$ over periods of length $k$. If $y$ has hump-shaped dynamics, the variance increases initially at a rate faster than $k$. Equivalently, $\text{Var}(y_t - y_{t-k})/k$ is hump shaped. Table 5 illustrates this pattern for (again) all maturities but the short rate. This perspective leads naturally to option volatilities, which are quoted on a similar per-period basis. Average annualized volatilities for interest rate caps and floors are pictured in Figure 6 for maturities between 1 and 10 years. While there are many reasons why implied volatilities might differ from conditional standard deviations of bond yields, the similarity is suggestive.

These five facts do not exhaust the known properties of bond yields, but they give us a place to start. Facts 1 and 3 describe the risk inherent in bond yields and the premium placed on this risk by the market. Fact 2 highlights the dynamic properties of bond yields, a critical feature for instruments that differ only in their maturities. The volatility and persistence of yields and spreads tells us indirectly that yields of different maturities are imperfectly correlated. Fact 4 suggests that interest rates are not normal, a feature of the data that could have a substantial impact on option prices. Fact 5 is a more subtle property of interest rate dynamics that has (as we will see) strong implications for model design.
3 Affine Models

Although the form varies, modern asset pricing theory is based on a single theoretical result: that in any arbitrage-free environment, there exists a positive random variable $m$ that satisfies

$$1 = E_t(m_{t+1}R_{t+1}),$$

(2)

for (one-period gross) returns $R$ on all traded assets at all dates $t$. We refer to $m$ as a pricing kernel. Bond pricing is a straightforward application. Since the one-period return on an $n + 1$-period bond is $R_{t+1} = b_{t+1}^{n}/b_{t}^{n+1}$, bond prices satisfy

$$b_{t+1}^{n+1} = E_t(m_{t+1}b_{t+1}^{n}).$$

(3)

Given a pricing kernel, we can compute bond prices recursively from the properties of $m$, starting with the initial condition $b_{t}^{0} = 1$ (a dollar today costs one dollar). Hence a bond pricing model consists of a description of $m$.

We explore the structure of the pricing kernel in affine models, in which yields of all maturities are linear functions of a vector of state variables. This class of models includes popular examples developed by, among others, Balduzzi, Das, and Foresi (1998), Balduzzi, Das, Foresi, and Sundaram (1996), Brennan and Schwartz (1979), Chen and Scott (1993), Cox, Ingersoll, and Ross (1985), Dai and Singleton (1997), Duffie and Kan (1996), Longstaff and Schwartz (1992), Pearson and Sun (1994), and Vasicek (1977). The ease of deriving implications for long maturities makes this class extremely attractive for applied work.

Our version of the class of affine models is taken from Duffie and Kan (1996), which we translate into discrete time. A vector of state variables $z$ follows

$$z_{t+1} - z_t = (I - \Phi)(\theta - z_t) + V(z_t)^{1/2}\varepsilon_t,$$

(4)

where $\{\varepsilon_t\} \sim \text{NID}(0, I)$, $V(z)$ is a diagonal matrix with typical element

$$v_i(z) = \alpha_i + \beta_i^\top z,$$

$\beta_i$ has nonnegative elements, and $\Phi$ is stable with positive diagonal elements. The pricing kernel is

$$-\log m_{t+1} = \delta + \gamma^\top z_t + \lambda^\top V(z_t)^{1/2}\varepsilon_t.$$

(5)

The process for $z$ requires that the volatility functions $v_i$ be positive. We define the set $D$ of admissible states as those values of $z$ for which volatility is nonnegative:

$$D = \{z : v_i(z) \geq 0 \text{ all } i\}.$$
Duffie and Kan (1996, Section 4) show (in the continuous-time analog) that \( z \) remains in \( D \) if the process satisfies

**Condition A**  For each \( i \) with \( \beta_i \neq 0 \):

(a) for all \( z \in D \) satisfying \( v_i(z) = 0 \) (the boundary of positive volatility), the drift is sufficiently positive: \( \beta_i^\top(I - \Phi)(\theta - z) > \beta_i^\top\beta_i/2 \); and

(b) if the \( j \)th component of \( \beta_i \) is nonzero for any \( j \neq i \) then \( v_i(z) \) and \( v_j(z) \) are proportional to each other (their ratio is a positive constant).

We refer to models characterized by (4, 5) and satisfying Condition A as the Duffie-Kan class of affine models.

The examples listed earlier are special cases of this general framework. In the Vasicek (1977) model, \( z \) is a scalar, \( \beta = 0, \gamma = 1, \) and \( \delta = \lambda^2 \alpha/2 \). In the Cox-Ingersoll-Ross (1985) model, \( z \) is also a scalar, \( \alpha = 0, \gamma = 1 + \lambda^2 \beta/2, \) and \( \delta = 0 \). In both cases, \( z \) is the short rate. Other examples will be described later.

Given this theoretical structure, bond prices are log-linear functions of the state:

\[
-\log b^n_t = A_n + B_n^\top z_t. \tag{6}
\]

The pricing relation (3) implies that the coefficients satisfy the difference equations

\[
A(n + 1) = A(n) + \delta + B(n)^\top(I - \Phi)\theta - \frac{1}{2} \sum_{j=1}^{k} (\lambda_j + B(n)_j)^2 \alpha_j \tag{7}
\]

\[
B(n + 1)^\top = (\gamma^\top + B(n)^\top \Phi) - \frac{1}{2} \sum_{j=1}^{k} (\lambda_j + B(n)_j)^2 \beta_j^\top, \tag{8}
\]

which are easily computed starting with the initial condition \( A(0) = 0 \) and \( B(0) = 0 \) (recall: \( b^n_0 = 1 \)).

Since yields and spreads are linear functions of the state, their moments follow from those of \( z \). The mean of \( z \) is \( \theta \). The covariance matrix for \( z \),

\[
\Gamma_0 \equiv E \left[ (z_t - \theta)(z_t - \theta)^\top \right],
\]

satisfies

\[
vec(\Gamma_0) = (I - \Phi \otimes \Phi)^{-1} vec[V(\theta)],
\]

\( 6 \)
where \( \text{vec}(A) \) is the vector formed from the columns of the matrix \( A \). Autocovariance matrices,
\[
\Gamma_k \equiv E \left[ (z_t - \theta)(z_{t-k} - \theta)^\top \right],
\]

obey
\[
\Gamma_{k+1} = \Phi \Gamma_k,
\]
for \( k \geq 0 \). See, for example, Harvey (1989, ch 8).

Yields and yield spreads are linear combinations of \( z \), so their properties can be computed from those of \( z \). Yields, for example, are
\[
y^n_t = n^{-1} \left( A_n + B_n^\top z_t \right),
\]
so their means are
\[
E(y^n_t) = n^{-1} \left( A_n + B_n^\top \theta \right).
\]
The variance of an arbitrary linear combination \( x_t = c^\top z_t \) is \( \text{Var}(x) = c^\top \Gamma_0 c \). Similarly, the \( k \)th autocovariance is \( \text{Cov}(x_t, x_{t-k}) = c^\top \Gamma_k c \) and the \( k \)th autocorrelation of \( x \) is
\[
\text{Corr}(x_t, x_{t-1}) = \frac{c^\top \Gamma_k c}{c^\top \Gamma_0 c}.
\]
Properties of multiperiod changes follow similar logic. Consider \( x_t - x_{t-k} = c^\top (z_t - z_{t-k}) \). Since
\[
\text{Var}(z_t - z_{t-k}) = 2\Gamma_0 - (\Gamma_k + \Gamma_k^\top) \equiv \Gamma^{(k)}
\]
\[
\text{Var}(x_t - x_{t-k}) = c^\top \Gamma^{(k)} c.
\]

These relations allow us to compute means, variances, and autocorrelations of yields and spreads from those of \( z \). In practice, of course, we do the reverse: find values of a model’s parameters that reproduce sample moments of bond yields.

4 Design Issues

The one-factor Cox-Ingersoll-Ross model provides a useful context for thinking about the properties of bond yields documented in Section 2. It’s grossly inadequate, but its weaknesses highlight changes we might want to make in designing better models. The model consists of the equations,
\[
\begin{align*}
  z_{t+1} - z_t &= (1 - \varphi) (\theta - z_t) + \sigma z_t^{1/2} \varepsilon_{t+1} \\
  -\log m_{t+1} &= \left[ 1 + (\lambda \sigma)^2 / 2 \right] z_t + \lambda \sigma z_t^{1/2} \varepsilon_{t+1},
\end{align*}
\]

(9) (10)
where \( \{\varepsilon_t\} \sim \text{NID}(0, 1) \). This is a special case of (4.5) with \( \alpha = \delta = 0, \beta = \sigma^2 \), and \( \gamma = 1 + (\lambda \sigma)^2 / 2 \) (a normalization that sets the short rate equal to \( z \)). Given this one-factor structure, yields and yield spreads are linear functions of the state variable \( z \). Since \( z \) is the short rate, we could estimate its mean, variance, and autocorrelation from sample moments of the short rate such as those reported in Table 1. Alternatively, we could use sample moments of yields of other maturities or of yield spreads.

Consider the ability of this model to account for the facts outlined earlier. Fact 1 concerns the slope and shape of the mean yield curve. The slope can be reproduced by choosing \( \lambda \) appropriately, but the shape is more difficult. Gibbons and Ramaswamy (1993) showed that with values of \( \varphi \) similar to observed autocorrelations of the short rate, the average yield curve has far less curvature than we see in the data.

Fact 2 is persistence. The one-factor structure means that every linear combination of bond yields has the same autocorrelation. Although the value of \( \varphi \) might be chosen to reproduce the autocorrelation of a particular bond yield, it cannot account for differences in persistence across maturities or between yields and spreads.

Fact 3 is conditional volatility. The (one-period) conditional variance in this model is linear in the state:

\[
Var_t(z_{t+1}) = \sigma^2 z_t.
\]

Since the short rate equals \( z \) in this model, there is an exact linear relation between the short rate and its conditional variance. Thus the model exhibits persistence in the conditional variance, but it cannot deliver the imperfect relation between the conditional variance and the short rate we see in the data (Figure 4).

Fact 4 is leptokurtosis. Leptokurtosis in \( \varepsilon \) is absent by construction (we assumed \( \varepsilon \) was normal). Moreover, the variation in the conditional variance is insufficient to generate as much kurtosis in yield changes as we see in the data. Consider the kurtosis of changes in \( z \). Since \( \varphi \) is close to one (see the autocorrelations of Table 1), changes in \( z \) are approximately

\[
z_t - z_{t-1} \approx \sigma z_{t-1}^{1/2} \varepsilon_t.
\]

The moments of changes in \( z \) include

\[
\begin{align*}
\text{Variance} & \cong E(z) \\
\text{Kurtosis} & \cong 3[E(z)^2 + Var(z)],
\end{align*}
\]
Excess kurtosis is

\[
\frac{\text{Kurtosis}}{\text{Variance}^2} = 3 \approx 3 \frac{\text{Var}(z)}{E(z)^2}.
\]

Since \( z = r \), we can get a rough idea of the magnitudes involved from the sample moments of \( y^1 \) in Table 1. Estimated excess kurtosis is (approximately) \( 3(2.699/6.683)^2 = 0.49 \), a factor of 20 smaller than the numbers reported in Tables 2 and 3.

Fact 5 is hump-shaped dynamics. In the model, yields and spreads inherit the properties of \( z \). All are AR(1)'s and thus have monotonic impulse response functions. Variances of changes are related to those of \( z \). The per-period variance, analogous to what we reported in Table 5, is

\[
\frac{\text{Var}(z_t - z_{t-k})}{k} = \frac{2(1 - \varphi^k)}{k} \text{Var}(z),
\]

which declines monotonically with the time interval \( k \). In short, the model is incapable of generating hump-shaped dynamics.

Each of these weaknesses points to possible solutions. For Fact 1 the standard solution is a multi-factor model. In one-factor settings, \( \varphi \) controls both the persistence of bond yields and the shape of the mean yield curve. With two or more factors, these roles can be separated. Multiple factors also help with Fact 2 (differences in persistence of the individual factors can produce differences in the persistence of yields and spreads) and Fact 3 (imperfect correlation of yields and their conditional variances). Fact 4 requires more direct intervention: non-normal behavior of one or more of the components of \( \varepsilon \). This moves us outside the Duffie-Kan class of models, but this can be done without abandoning the convenient log-linearity between bond prices and state variables (Das and Foresi 1996; Duffie, Pan, and Singleton 1999). Fact 5 is the most difficult. Since affine models with multiple independent factors have monotonic dynamics, we clearly have to move beyond them. The question is which direction.

5 Reasonable Models

The following two-factor model is probably the simplest example that approximates all of the facts documented earlier:

\[ z_{1t+1} - z_{1t} = (1 - \varphi_{11})(\theta_2 - z_{1t}) - \varphi_{12}(\theta_2 - z_{2t}) + \sigma_1 \varepsilon_{1t+1} \] \hspace{1cm} (11)

\[ z_{2t+1} - z_{2t} = (1 - \varphi_{22})(\theta_2 - z_{2t}) + \sigma_2 z_{2t}^{1/2} \varepsilon_{2t+1} \] \hspace{1cm} (12)

\[ -\log m_{t+1} = \delta + z_{3t} + [1 + (\lambda_2 \sigma_2)^2/2] z_{2t} + \lambda_1 \sigma_1 \varepsilon_{1t+1} + \lambda_2 \sigma_2 z_{2t}^{1/2} \varepsilon_{2t+1}, \] \hspace{1cm} (13)
where $\delta = (\lambda_1 \sigma_1)^2 \log E(e^{\varepsilon_1^2})$, $\varepsilon_1$ and $\varepsilon_2$ are independent iid sequences, $\{\varepsilon_{2t}\} \sim \text{NID}(0, 1)$, and $\{\varepsilon_{It}\}$ has mean zero, variance one, and arbitrary excess kurtosis $\kappa_1$. We refer to this as our preferred model. The key ingredients of these two factors, interaction between factors ($\varphi_{12}$), and the possibility of excess kurtosis ($\kappa_1$). The structure is a combination of a CIR-like square root factor ($z_2$) and a Vasicek-like factor ($z_1$) with a some non-normality sprinkled in. We’ll describe the logic behind it shortly.

We estimate the model by translating the facts into moment conditions and applying GMM. [Standard errors to come.] We use 9 moments to estimate the 9 parameters of the model: the mean of the short rate, the mean of the 6-month and 10-year spreads (slope and shape of yield curve), the autocorrelation of the 6-month rate and the 6-month spread (persistence), the variance of 1- and 3-period changes in the 6-month rate (the hump) and 1-period changes in the 6-month spread, and the kurtosis of one-period changes in the 6-month rate (kurtosis). (Our focus on the 6-month rate mirrors the emphasis on 6-month LIBOR in fixed income derivatives markets.) Since there are equal numbers of moments and parameters, the model reproduces these features of the data exactly. The results of this exercise are reported in Table 6 as Model D.

The easiest way to understand the structure of the preferred model is to consider weaknesses in simpler models. Consider a two-factor Cox-Ingersoll-Ross model. In the notation of the general affine model [equations (4,5)], $z$ is two-dimensional, $\alpha_0 = 0$, $\beta_1, \beta_2 = \text{diag}(\sigma_1^2, \sigma_2^2)$, $\Phi = \text{diag}(\varphi_{11}, \varphi_{22})$, $\gamma_1 = 1 + (\lambda_1 \sigma_1)^2/2$, and $\delta = 0$. These choices produce a short rate of $z_{1t} + z_{2t}$. Parameter estimates are reported in Table 6 as Model A. The estimates include the arbitrary restriction $\theta_1 = \theta_2$, since it is notoriously difficult to estimate them separately. They also ignore two moments: the variance of 3-period changes and the kurtosis of 1-period changes in the 6-month rate. [To come: report J-statistic.]

The second factor is a significant improvement. We ask the model to reproduce both the strong autocorrelation of the 6-month rate (0.970, Table 1) and the more modest autocorrelation of the 6-month spread (0.680, Table 2). The estimates accomplish this by combining the effects of a persistent factor $\varphi_{22} = 0.980$ and a less persistent one $\varphi_{11} = 0.515$. The curvature of the mean yield curve then comes from placing a greater “price of risk” on the first factor than the second (note that $\lambda_1$ is larger in absolute value than $\lambda_2$). The model fails, however, to explain two properties of the data noted earlier: interest rates exhibit monotonic dynamics and very little kurtosis (excess kurtosis of 1-period changes in the 6-month rate is about 1.2,
far less than we see in the data). The reasoning is similar to that of the one-factor model.

Model B introduces an off-diagonal element \( \varphi_{12} \) into the two-factor CIR model:

\[
\begin{align*}
\zeta_{t+1} - \zeta_t &= (1 - \varphi_{11})(\theta_2 - \zeta_t) - \varphi_{12}(\theta_2 - \zeta_2t) + \sigma_1 \zeta_{1t}^{1/2} \varepsilon_{1t+1} \\
\zeta_{2t+1} - \zeta_{2t} &= (1 - \varphi_{22})(\theta_2 - \zeta_{2t}) + \sigma_2 \zeta_{2t}^{1/2} \varepsilon_{2t+1} \\
- \log m_{t+1} &= \sum_{i=1}^{2} \left( [1 + (\lambda_i \sigma_i)^2/2] \varepsilon_{it} + \lambda_i \sigma_i \zeta_{it}^{1/2} \varepsilon_{it+1} \right),
\end{align*}
\]

If \( \varphi_{12} = 0 \) this reduces to the two-factor CIR model. If \( \varphi_{12} = 1 - \varphi_{11} \) this has the flavor of the central tendency model proposed by Balduzzi, Das, and Foresi (1998). The law of motion for \( \zeta_1 \) becomes

\[
\zeta_{1t+1} - \zeta_{1t} = (1 - \varphi_{11})(\theta_2 - \zeta_{1t}) + \sigma_1 \varepsilon_{1t+1},
\]

approaching the “central tendency” \( \zeta_2 \) at a rate governed by \( 1 - \varphi_{11} \). More generally, nonzero values of \( \varphi_{12} \) allow for the possibility of nonmonotonic dynamics. The estimated value is larger than the central tendency in the sense that \( \varphi_{12} > 1 - \varphi_{11} \).

We choose \( \varphi_{12} \) to reproduce the variance of 3-period changes in the 6-month rate (the hump in volatility documented in Table 5), so the “triangular \( \Phi \)” model is clearly capable of generating hump-shaped dynamics. Unfortunately, it violates Condition A(a), the multivariate analog of the Feller condition. The condition tells us that at \( \zeta_1 = 0 \), the following condition must hold for all nonnegative values of \( \zeta_2 \):

\[
(1 - \varphi_{11}) \theta_1 - \varphi_{12}(\theta_2 - \zeta_2) \geq 0.
\]

Evidently \( \varphi_{12} \) must be positive; otherwise large values of \( \zeta_2 \) would violate the condition. There’s also a limit to how large it can be. If \( \zeta_2 = 0 \) (it’s lower bound), the condition becomes, \( (1 - \varphi_{11}) \theta_1 - \varphi_{12} \theta_2 \geq 0 \). With \( \theta_1 = \theta_2 \), as above, \( 1 \geq \varphi_{11} + \varphi_{12} \), which is violated by the estimates in the table. Stated simply: the data require a large value of \( \varphi_{12} \), but with a value as large as we estimate, we can no longer guarantee that \( \zeta_1 \) remains positive. This isn’t a theorem, but in practice we have found it difficult to reconcile Condition A(a) with hump-shaped dynamics when both state variables are required to be positive.

That leads us to Model C, in which we have eliminated the square root in the innovation term of equation (14). The result is the preferred model, equations (11,12,13), without excess kurtosis in \( \varepsilon_1 \) (ie, \( \kappa_1 = 0 \)). We can no longer guarantee that interest rates are positive, but like others before us we find the convenience
of a linear model enough to forego this requirement. [Talk about how we treat the means.] We now have a model that generates hump-shaped dynamics. But with normal innovations, the kurtosis of yield changes is still slight. [Number?]

Our preferred model is D, which introduces the possibility of excess kurtosis to Model C. We put kurtosis in the first factor only, both because it’s more difficult to put into positive factors like $z_2$ and because it allows us to match the the general features of the data.

To summarize: the properties of bond yields call for a model in which there is interaction among the factors (a role played in our model by $\varphi_{12}$) and excess kurtosis of yield changes ($\kappa_1$). Other research suggests that even a two-factor model is insufficient, and points to further extensions....

6 Interest Rate Caps

<incomplete>

To this point we have focused exclusively on data from the U.S. treasury market in guiding the process of design and estimation. Given our emphasis on various patterns of volatility, it seems natural to exploit information from option markets as well. In this section we ask what features of affine models are important in accounting for the average shape of the term structure of implied volatilities documented in Figure 6.

What one might expect is as follows. The average shape of the term structure of implied volatilities will be primarily affected by two forces. The first is the dynamics of the model (especially $\varphi_{12}$): the per-period conditional variance of future bond yields is hump-shaped. The second is excess kurtosis ($\kappa_1$). Although kurtosis increases the value of out-of-the-money options, it reduces the value of at-the-money options. Moreover, the effect is greater for short options: the Central Limit Theorem means that the excess kurtosis falls with the length of the time horizon, so implied volatilities of short options fall more than those of long options.

An interest rate cap contract is a collection of call options written on simply compounded interest rates, with the same strike but different maturities. Each individual call is referred to as a caplet. If we denote the payment frequency in
years — often called the tenor of the cap contract — as \( \tau \) (e.g., \( \tau = 1/2 \) indicates semi-annual payments), then the payoff accruing to the \( j \)th caplet is

\[
\tau L \left( R_{t+j\tau}^r - K \right)^+, 
\]

where \( K \) is the strike, \( L \) is the notional principal, and \( R_{t+j\tau}^r \) is the simply compounded interest rate on a \( \tau \) year bond at time \( t + j\tau \). The simply compounded interest rate is related to the continuously compounded yield, \( y^r \), as follows.

\[
R^r = \frac{1}{\tau} \left( e^{\tau y^r} - 1 \right) .
\]

We focus on caps which are settled ‘in arrears,’ which means that the payment accruing to the \( j \)th caplet is received \((i + 1)\tau\) years in the future, \( \tau \) years after the magnitude of the payment is determined.

We denote \( t_1 = t + j\tau \) and \( t_2 = t_1 + \tau \). Given the above definition of simple interest, the value of the \( j \)th caplet is

\[
c^j_t = \tau L E_t \left[ m_{t_2}(R_{t+j\tau}^r - K)^+ \right] \\
= L E_t \left[ m_{t_1} \left( \exp(\tau y_{t_1}^r) - (1 + \tau K) \right)^+ \right] \\
= L E_t \left[ m_{t_1} E_{t_1} m_{t_2} \left( \frac{1}{b_{t_1}^r} - (1 + \tau K) \right)^+ \right] \\
= L E \left[ m_{t_1} (1 - (1 + \tau K)b_{t_1}^r)^+ \right] .
\tag{15}
\]

We follow Duffie, Pan and Singleton (1999), Heston (1993) and others in using transform methods to evaluate (15). We begin by defining \( G_{a,b,n}(y) \) as price of the payoff \( \exp(a^\top z_{t+n}) \) at time \( t + n \), conditional on the event \( b^\top z_{t+n} \leq y \):

\[
G_{a,b,n}(y) = E_t \left[ m_{t+n} e^{a^\top z_{t+n}} I_{b^\top z_{t+n} \leq y} \right] ,
\tag{16}
\]

where \( I \) is the indicator function, taking on a value of unity if \( b^\top z_{t+n} \leq y \). The function \( G \) has a number of natural interpretations. \( G_{0,1,n}(\infty) \) is the price of an \( n \)-period bond, \( G_{0,1,n}(y) \) is an Arrow-Debreu state price, and \( G_{a,b,n}(y) \) is an arbitrary contingent claim price. By combining equations (15) and (16) we can express the caplet price in terms of the function \( G \).

\[
c^j_t = L \times \left( G_{0,-B(\tau),n}(y) - (1 + \tau K) e^{-A(\tau)} G_{-B(\tau),-B(\tau),n}(y) \right) \\
y = -\ln(1 + \tau K) + A(\tau) ,
\tag{17}
\]

13
where, recall, the parameters $A(n), B(n)$ are the coefficients from the affine bond pricing function, equation (6).

In general, the expression (16) does not permit direct calculation of the function $G$. Instead, we obtain the Fourier transform of $G$ and then evaluate its inverse numerically. We use $\mathcal{F}$ to denote the transform:

$$
\mathcal{F}_{a,b,n} (\omega) = \int_{-\infty}^{+\infty} e^{i\omega y} dG_{a,b,n} (y) = E_t \left( m_{t+n} e^{i\omega (a^\top + ib^\top) z_{t+n}} \right).
$$

Thus, the Fourier transform of $G$ has the interpretation of the price of the complex-valued payoff, $\exp(\omega (a^\top + ib^\top) z_{t+n})$.

We use $\phi(u, z_t, n)$ to denote the time $t$ value of the date $t+n$ payoff $\exp(u^\top z_{t+n})$. It is straightforward to show that $\phi$ is exponential-affine in $z$:

$$
\phi(u, z_t, n) = E_t \left[ m_{t+n} \exp(-u^\top z_{t+n}) \right] = \exp \left( -C(n) - D(n)^\top z_t \right),
$$

where the coefficients $C(n)$ and $D(n)$ obey recursions identical to those governing the bond pricing coefficients, $A(n)$ and $B(n)$, (described in (6)), with the exception of the initial condition, $D(0) = -u$. The Fourier transform of $G_{a,b,n}(y)$ is therefore $\phi(a + ib\omega, z_t, n)$ which we know up to a set of recursively defined coefficients, $C(n)$ and $D(n)$.

The final step is to obtain $G$ by numerically integrating the following inversion of $\mathcal{F}$:

$$
G_{a,b,n} (y) = \frac{\phi(a)}{2} + \frac{1}{\pi} \int_0^\infty \frac{\text{Re} \left[ \phi(a + ib\omega) \right] \sin(\omega y) - \text{Im} \left[ \phi(a + ib\omega) \right] \cos(\omega y)}{\omega} d\omega.
$$

Given values for $G$ we obtain caplet prices via (15) and then sum them up to arrive at the value of a specific cap contract.

**6.1 Implied Volatility from Affine Models**

<to be added>
7 Final Thoughts

In an important paper, Dai and Singleton (1997) characterize the class of $N$ factor affine term structure models as consisting of $N + 1$ sub-classes. They go on to estimate a maximal model in each class and discuss the various strengths and weaknesses of affine models in this context. Our approach goes in the opposite direction. We start with a list of what we view as the salient features of data on U.S. interest rates and then attempt to find a parsimonious affine specification which accounts for them. What we gain are some helpful insights into the mapping between parameters and data. The parameters which govern stochastic volatility, for instance, do not seem important in accounting for the striking patterns in the kurtosis of changes in yields which we observe in data. Parameters which govern the correlation between factors, on the other hand, turn out to be quite important in accounting for both the dynamics in interest rates — the hump shaped impulse response functions — as well as the various hump shapes observed in volatility, both historic and implied, as the time horizon grows.
References


Table 1
Properties of US Government Bond Yields

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Mean</th>
<th>St Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Auto</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>6.683</td>
<td>2.699</td>
<td>1.073</td>
<td>1.256</td>
<td>0.959</td>
</tr>
<tr>
<td>3 months</td>
<td>7.039</td>
<td>2.776</td>
<td>1.007</td>
<td>0.974</td>
<td>0.971</td>
</tr>
<tr>
<td>6 months</td>
<td>7.297</td>
<td>2.769</td>
<td>0.957</td>
<td>0.821</td>
<td>0.971</td>
</tr>
<tr>
<td>9 months</td>
<td>7.441</td>
<td>2.721</td>
<td>0.924</td>
<td>0.733</td>
<td>0.970</td>
</tr>
<tr>
<td>12 months</td>
<td>7.544</td>
<td>2.667</td>
<td>0.903</td>
<td>0.669</td>
<td>0.970</td>
</tr>
<tr>
<td>24 months</td>
<td>7.819</td>
<td>2.491</td>
<td>0.887</td>
<td>0.484</td>
<td>0.973</td>
</tr>
<tr>
<td>36 months</td>
<td>8.008</td>
<td>2.370</td>
<td>0.913</td>
<td>0.378</td>
<td>0.976</td>
</tr>
<tr>
<td>48 months</td>
<td>8.148</td>
<td>2.284</td>
<td>0.944</td>
<td>0.321</td>
<td>0.977</td>
</tr>
<tr>
<td>60 months</td>
<td>8.253</td>
<td>2.221</td>
<td>0.971</td>
<td>0.288</td>
<td>0.978</td>
</tr>
<tr>
<td>84 months</td>
<td>8.398</td>
<td>2.138</td>
<td>1.008</td>
<td>0.251</td>
<td>0.979</td>
</tr>
<tr>
<td>120 months</td>
<td>8.529</td>
<td>2.069</td>
<td>1.035</td>
<td>0.217</td>
<td>0.981</td>
</tr>
</tbody>
</table>

The data are end-of-month estimates of continuously-compounded zero-coupon US government bond yields expressed as annual percentages. They were supplied by Robert Bliss (“smoothed Fama-Bliss” method) and cover the period January 1970 to December 1995 (312 observations). Mean is the sample mean, St Dev the sample standard deviation, Skewness an estimate of the skewness measure γ₁, Kurtosis an estimate of the (excess) kurtosis measure γ₂, and Auto an estimate of the first autocorrelation. The skewness and kurtosis measures are defined in terms of central moments μᵢ: γ₁ = μ₃/μ₂³/² and γ₂ = μ₄/μ₂⁴ - 3. Both are zero for normal random variables. Our estimates replace population moments with sample moments: the estimate of μᵢ for random variable y and moment j > 1 is \( \frac{\sum_{t=1}^{T} (y_t - \bar{y})^j}{T} \), where \( \bar{y} = \frac{\sum_{t=1}^{T} y_t}{T} \) is the sample mean.
Table 2  
Properties of Yield Spreads and Monthly Changes in Yields

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Mean</th>
<th>St Dev</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Auto</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 months</td>
<td>0.357</td>
<td>0.350</td>
<td>1.811</td>
<td>5.642</td>
<td>0.403</td>
</tr>
<tr>
<td>6 months</td>
<td>0.614</td>
<td>0.507</td>
<td>1.325</td>
<td>4.538</td>
<td>0.527</td>
</tr>
<tr>
<td>9 months</td>
<td>0.758</td>
<td>0.601</td>
<td>1.117</td>
<td>4.647</td>
<td>0.618</td>
</tr>
<tr>
<td>12 months</td>
<td>0.861</td>
<td>0.683</td>
<td>0.863</td>
<td>4.177</td>
<td>0.680</td>
</tr>
<tr>
<td>24 months</td>
<td>1.136</td>
<td>0.929</td>
<td>-0.017</td>
<td>2.197</td>
<td>0.792</td>
</tr>
<tr>
<td>36 months</td>
<td>1.325</td>
<td>1.095</td>
<td>-0.424</td>
<td>1.368</td>
<td>0.833</td>
</tr>
<tr>
<td>48 months</td>
<td>1.465</td>
<td>1.213</td>
<td>-0.599</td>
<td>0.983</td>
<td>0.855</td>
</tr>
<tr>
<td>60 months</td>
<td>1.570</td>
<td>1.300</td>
<td>-0.677</td>
<td>0.765</td>
<td>0.868</td>
</tr>
<tr>
<td>84 months</td>
<td>1.715</td>
<td>1.420</td>
<td>-0.728</td>
<td>0.532</td>
<td>0.882</td>
</tr>
<tr>
<td>120 months</td>
<td>1.846</td>
<td>1.526</td>
<td>-0.735</td>
<td>0.385</td>
<td>0.892</td>
</tr>
</tbody>
</table>

A. Spreads Over the Short Rate: $y_t^n - y_t^1$

B. Monthly Changes in Yields: $y_{t+1}^n - y_t^n$

1 month     | -0.009 | 0.764  | -1.043   | 9.939    | 0.003|
3 months    | -0.010 | 0.662  | -1.747   | 11.158   | 0.146|
6 months    | -0.010 | 0.656  | -1.352   | 10.691   | 0.150|
9 months    | -0.009 | 0.648  | -1.088   | 10.514   | 0.153|
12 months   | -0.009 | 0.634  | -0.926   | 10.251   | 0.154|
24 months   | -0.009 | 0.558  | -0.583   | 7.902    | 0.157|
36 months   | -0.009 | 0.499  | -0.424   | 5.829    | 0.157|
48 months   | -0.009 | 0.461  | -0.345   | 4.598    | 0.152|
60 months   | -0.008 | 0.434  | -0.300   | 3.845    | 0.143|
84 months   | -0.008 | 0.401  | -0.256   | 3.000    | 0.120|
120 months  | -0.007 | 0.374  | -0.220   | 2.337    | 0.094|

See Table 1 for definitions and data sources.
Table 3
Higher Moments of Yield “Residuals”

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Raw Residuals</th>
<th>Normalized Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skewness</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>1 month</td>
<td>-0.397</td>
<td>9.287</td>
</tr>
<tr>
<td>3 months</td>
<td>-1.160</td>
<td>10.761</td>
</tr>
<tr>
<td>6 months</td>
<td>-0.678</td>
<td>9.991</td>
</tr>
<tr>
<td>9 months</td>
<td>-0.378</td>
<td>9.249</td>
</tr>
<tr>
<td>12 months</td>
<td>-0.216</td>
<td>8.509</td>
</tr>
<tr>
<td>24 months</td>
<td>-0.008</td>
<td>5.932</td>
</tr>
<tr>
<td>36 months</td>
<td>0.042</td>
<td>4.230</td>
</tr>
<tr>
<td>48 months</td>
<td>0.060</td>
<td>3.312</td>
</tr>
<tr>
<td>60 months</td>
<td>0.070</td>
<td>2.797</td>
</tr>
<tr>
<td>84 months</td>
<td>0.078</td>
<td>2.283</td>
</tr>
<tr>
<td>120 months</td>
<td>0.084</td>
<td>1.881</td>
</tr>
</tbody>
</table>

Entries are estimates of skewness and excess kurtosis based on residuals $\eta$ from an AR(3) for bond yields of the relevant various maturity. Raw Residuals refers to these residuals, $\eta$. Normalized Residuals refers to $\eta_t h_{t-1}^{1/2}$, where $h$ is computed recursively from

$$h_t = \varphi h_{t-1} + (1 - \varphi)\eta_t^2,$$

with $\varphi = 0.95$, starting with $h_0$ equal to the estimated unconditional variance of $\eta$ (its standard deviation squared).
Table 4
ARMA Representations of Bond Yields

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Schwartz Criterion</th>
<th>Akaike Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best Model</td>
<td>Hump?</td>
</tr>
<tr>
<td>1 month</td>
<td>ARMA(1,0)</td>
<td>no</td>
</tr>
<tr>
<td>3 months</td>
<td>ARMA(1,1)</td>
<td>yes</td>
</tr>
<tr>
<td>6 months</td>
<td>ARMA(1,1)</td>
<td>yes</td>
</tr>
<tr>
<td>9 months</td>
<td>ARMA(1,1)</td>
<td>yes</td>
</tr>
<tr>
<td>12 months</td>
<td>ARMA(1,1)</td>
<td>yes</td>
</tr>
<tr>
<td>24 months</td>
<td>ARMA(1,1)</td>
<td>yes</td>
</tr>
<tr>
<td>36 months</td>
<td>ARMA(1,1)</td>
<td>yes</td>
</tr>
<tr>
<td>48 months</td>
<td>ARMA(2,0)</td>
<td>yes</td>
</tr>
<tr>
<td>60 months</td>
<td>ARMA(2,0)</td>
<td>yes</td>
</tr>
<tr>
<td>84 months</td>
<td>ARMA(2,0)</td>
<td>yes</td>
</tr>
<tr>
<td>120 months</td>
<td>ARMA(2,0)</td>
<td>yes</td>
</tr>
</tbody>
</table>

The table reports best ARMA models for bond yields — with best defined by the Schwartz and Akaike criteria, respectively — and whether the impulse response function of the estimated best model exhibits a “hump.” Examples are pictured in Figure 5.
Table 5
Standard Deviations of Multiperiod Changes in Yields

<table>
<thead>
<tr>
<th>Maturity n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>0.772</td>
<td>0.772</td>
<td>0.751</td>
<td>0.692</td>
<td>0.644</td>
</tr>
<tr>
<td>3 months</td>
<td>0.668</td>
<td>0.715</td>
<td>0.714</td>
<td>0.680</td>
<td>0.637</td>
</tr>
<tr>
<td>6 months</td>
<td>0.660</td>
<td>0.707</td>
<td>0.707</td>
<td>0.663</td>
<td>0.617</td>
</tr>
<tr>
<td>9 months</td>
<td>0.652</td>
<td>0.699</td>
<td>0.695</td>
<td>0.643</td>
<td>0.598</td>
</tr>
<tr>
<td>12 months</td>
<td>0.637</td>
<td>0.684</td>
<td>0.677</td>
<td>0.624</td>
<td>0.581</td>
</tr>
<tr>
<td>24 months</td>
<td>0.559</td>
<td>0.602</td>
<td>0.596</td>
<td>0.552</td>
<td>0.523</td>
</tr>
<tr>
<td>36 months</td>
<td>0.500</td>
<td>0.539</td>
<td>0.535</td>
<td>0.499</td>
<td>0.483</td>
</tr>
<tr>
<td>48 months</td>
<td>0.461</td>
<td>0.496</td>
<td>0.493</td>
<td>0.462</td>
<td>0.456</td>
</tr>
<tr>
<td>60 months</td>
<td>0.435</td>
<td>0.465</td>
<td>0.463</td>
<td>0.437</td>
<td>0.437</td>
</tr>
<tr>
<td>84 months</td>
<td>0.401</td>
<td>0.425</td>
<td>0.423</td>
<td>0.404</td>
<td>0.411</td>
</tr>
<tr>
<td>120 months</td>
<td>0.374</td>
<td>0.391</td>
<td>0.389</td>
<td>0.376</td>
<td>0.391</td>
</tr>
</tbody>
</table>

Entries are “per month” standard deviations of changes in bond yields over different time intervals. For yields $y$ of a given maturity, we computed the sample standard deviation of $y_t - y_{t-k}$ over the period January 1971 to December 1995, using the initial year of data for lags. In the table, the standard deviations have been divided by $k^{1/2}$ to put them on a per period basis.
### Table 6
Estimated Parameters of Two-Factor Models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>0.0027845</td>
<td>0.0027845</td>
<td>0.0027845</td>
<td>$\theta_2$</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.0027845</td>
<td>0.0027845</td>
<td>0.0027845</td>
<td>0.0055690</td>
</tr>
<tr>
<td>$\varphi_{11}$</td>
<td>0.51539</td>
<td>0.53431</td>
<td>0.49387</td>
<td>0.49387</td>
</tr>
<tr>
<td>$\varphi_{12}$</td>
<td>0.0</td>
<td>0.78611</td>
<td>0.76855</td>
<td>0.76855</td>
</tr>
<tr>
<td>$\varphi_{22}$</td>
<td>0.98035</td>
<td>0.95879</td>
<td>0.95971</td>
<td>0.95971</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.011808</td>
<td>0.0081447</td>
<td>0.018922</td>
<td>0.018922</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>0.008828</td>
<td>0.0045808</td>
<td>0.0052384</td>
<td>0.0052384</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-849.23</td>
<td>-1289</td>
<td>-2436.8</td>
<td>-2436.8</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-103.36</td>
<td>-282.29</td>
<td>-251.7</td>
<td>-251.7</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>9.83</td>
</tr>
</tbody>
</table>

Entries are estimates of the parameters of these models:

- Model A: 2-factor CIR.
- Model B: triangular $\Phi$ model (2-factor CIR plus nonzero value of $\varphi_{12}$).
- Model C: preferred model with $\kappa_1 = 0$ (no excess kurtosis).
- Model D: preferred model.
Figure 1
Autocorrelation Functions for Yields and Spreads

The lines are autocorrelation functions for the 1-month yield ($y^1$, solid line), the 5-year yield ($y^0$, dashed line), and the 2-year spread ($y^{21} - y^1$, dash-dotted line). The data are described in the notes to Table 1.
Figure 2
Autocorrelation Functions for Squared Residuals

The lines are indications of persistence in volatility. Both are based on residuals from a regression of bond yields on three lags. The autocorrelation functions pertain to squared residuals for the 1-month yield ($y^1$, solid line), the 1-year yield ($y^{12}$, dashed line), and the 5-year yield ($y^{50}$, dash-dotted line). The data are described in the notes to Table 1.
Figure 3
Estimated Volatilities of Bond Yields

The lines are estimates of volatility for three maturities: the 1-month yield ($y^1$, solid line), the 1-year yield ($y^{12}$, dashed line), and the 5-year yield ($y^{50}$, dash-dotted line). They are constructed from residuals $\eta$ of bond yields on three lags using the “RiskMetrics” update rule:

$$h_t = \varphi h_{t-1} + (1 - \varphi) \eta_t^2,$$

with $\varphi = 0.95$. The initial value is the sample variance of $\eta$. The numbers in the figure are $h_t^{1/2}$ divided by the sample standard deviation.
Figure 4
Short Rate Volatility: Relation to Short Rate

The figure is a scatterplot of estimated conditional variance ($h$) against the level of the short rate ($y^1$). Data are monthly, April 1970 to December 1995.
Figure 5
Impulse Response Functions for Bond Yields

The lines are moving average coefficients implied by ARMA models in Table 4 (Schwartz criterion). They refer to models of the 1-month yield ($y^1$, solid line), the 1-year yield ($y^{12}$, dashed line), and the 5-year yield ($y^{60}$, dash-dotted line).
Figure 6
Mean Volatilities for Caps and Floors

Lines are mean volatilities of semiannual interest rate caps and floors with maturities between 1 and 10 years. The data are implied volatilities (Black’s formula applied to the 6-month rate) for at-the-money interest rate caps and floors, expressed as annual percentages. They were supplied by Datastream and cover the period February 1, 1995 to March 24, 1998 (820 daily observations).