PREDICTABLE CHANGES IN YIELDS AND FORWARD RATES

David Backus
Silverio Foresi
Abon Mozumdar
Liuren Wu

Working Paper **6379**

NBER WORKING PAPER SERIES

PREDICTABLE CHANGES IN YIELDS AND FORWARD RATES

David Backus
Silverio Foresi
Abon Mozumdar
Liuren Wu

Working Paper 6379
http://www.nber.org/papers/w6379

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 1998

We thank Robert Bliss for providing the computer programs and data used to estimate forward rates and yields; Darrell Duffie for advice on the Feller condition; Allan Gregory and Stanley Zin for insight into asymptotic distribution theory; Sanford Grossman for the suggestion to explore the implications for trading strategies; Robert Hodrick and other participants at the Spring 1997 meeting of the NBER's Asset Pricing group for comments on earlier drafts; and a referee for unusually constructive suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect those of Salomon Smith Barney or the National Bureau of Economic Research. The most recent version of this paper is available at http://www.stern.nyu.edu/~dbackus.

Predictable Changes in Yields and Forward Rates
David Backus, Silverio Foresi, Abon Mozumdar
and Liuren Wu
NBER Working Paper No. 6379
January 1998
JEL Nos. E43, G12

ABSTRACT

We consider patterns in the predictability of interest rates, characterized relative to the expectations hypothesis (EH), and attempt to account for them with affine models. We make the following points: (i) Discrepancies in the data from the EH take a particularly simple form with forward rates: as theory suggests, the largest discrepancies are at short maturities. (ii) Reasonable estimates of one-factor Cox-Ingersoll-Ross models imply regressions on the opposite side of the EH than we see in the data: regression slopes are greater than one, not less than one. (iii) Multifactor affine models can nevertheless approximate both departures from the EH and other properties of interest rates.

David Backus
Stern School of Business
New York University
44 West 4th Street
New York, NY 10012-1126
and NBER
dbackus@stern.nyu.edu

Silverio Foresi
Salomon Smith Barney
7 World Trade Center
New York, NY 10048
sforesi@sbi.com

Abon Mozumdar
Department of Finance
Virginia Tech University
Blacksburg, VA 24061-0221
abon@vt.edu

Liuren Wu
Stern School of Business
New York University
44 West 4th Street
New York, NY 10012-1126
lwu1@stern.nyu.edu

1 Introduction

Although textbooks continue to pay lip service to the expectations hypothesis of the term structure of interest rates, experts have long regarded it a poor approximation of the evidence. Most conclude, instead, that long interest rates contain both forecasts of future interest rates and term premiums that vary through time. An enormous body of research to this effect has been surveyed repeatedly, most recently by Bekaert, Hodrick, and Marshall (1996), Campbell (1995), and Evans and Lewis (1994). The most common evidence against the expectations hypothesis involves regressions of future yields or forward rates on current interest rate spreads. Postwar US data suggests that while current spreads are useful predictors of future interest rates, the form of this prediction differs substantially from the expectations hypothesis.

The weight of the evidence motivates a search for models that might account for observed patterns of predictability in interest rates. We study a subset of the affine yield models characterized by Duffie and Kan (1996). Earlier work by Frachot and Lesne (1994) documented the ability of the one-factor Cox-Ingersoll-Ross (1985) model to account for observed departures from the expectations hypothesis with term premiums that vary through time. We go farther in constructing models that account for both predictable variation in term premiums and other salient features of interest rates.

We make four contributions that concern both the nature of the evidence and its theoretical explanation. The first is to cast the evidence in a new form based on the martingale property of forward rates under the expectations hypothesis. In this form, the largest differences from the expectations hypothesis are for maturities under two years. This feature of the data conforms with a broad class of stationary bond pricing models, but contrasts with more popular tests of the expectations hypothesis based on yields (Bekaert, Hodrick, and Marshall 1996, Campbell 1995, and Evans and Lewis 1994, for example). The second is to illustrate the difficulty of accounting for this evidence with the one-factor Cox-Ingersoll-Ross model. The one-factor model cannot account simultaneously for deviations from the expectations hypothesis and the average upward slope of the yield curve. The third is to propose and estimate a new model in the affine class in which the short rate depends negatively on one or more "square-root" factors. We use this model to account for both the evidence against the expectations hypothesis and other features of interest rates. We argue that they approximate the data better than the two-factor Cox-Ingersoll-Ross models studied by Roberds and Whiteman (1996). The fourth is to compare our forward rate regressions (in which the largest deviations from the expectations hypothesis are at short maturities) to yield regressions (in which the largest deviations are at long maturities). We suggest that the large numerical differences between these two sets of regressions masks a fundamental similarity in their information content.

2 Notation and Data

In what follows, the continuously-compounded yield on an n-period bond at date t is denoted y_t^n and defined by

$$y_t^n = -n^{-1} \log b_t^n, \tag{1}$$

where b_t^n is the dollar price at date t of a claim to one dollar at $t+n$. Forward rates are defined by

$$f_t^n = \log(b_t^n / b_t^{n+1}), \tag{2}$$

so that yields are averages of forward rates:

$$y_t^n = n^{-1} \sum_{i=0}^{n-1} f_t^i. \tag{3}$$

The short rate is $r_t = y_t^1 = f_t^0$.

In practice, yields and forward rates are estimated rather than observed. From prices of bonds for a variety of maturities, the "discount function" b_t^n (viewed as a function of n at each date t) is interpolated between missing maturities n and smoothed to reduce the impact of noise (nonsynchronous price quotes, bid/ask spreads, and so on). There is no generally-accepted best practice for doing this. We follow Bliss (1996) in using four methods: Smoothed Fama-Bliss, Unsmoothed Fama-Bliss, McCulloch Cubic Spline, and Extended Nelson-Siegel.

The estimation methods are described in detail by Bliss (1996). Fama and Bliss's unsmoothed method extracts forward rates from prices of bonds of successive maturities using a relation analogous to (2) for coupon bonds. McCulloch approximates the discount function with a cubic spline. Nelson and Siegel's extended method approximates the yield curve, rather than the discount function, with a smooth function that gives the long end of the yield curve a horizontal asymptote. Fama and Bliss's smoothed method applies a similar approximation to yields implied by their unsmoothed forward rates. Among the four methods, Unsmoothed Fama-Bliss sticks out in not smoothing the raw data in some way.

We apply all four methods to bond prices collected by the Center for Research in Securities Prices (CRSP) at the University of Chicago using computer programs and data and supplied by Bliss. Academic research (Elton and Green 1996) and Wall Street folklore suggest that CRSP bond prices have more noise in them than those used in industry, but data of higher quality are not available over sufficiently long sample periods. Elton and Green, for example, study just 3 years of intraday data. Our sample extends from January 1970 to December 1995 (312 monthly observations).

3 Forward Rate Regressions

We use these four data sets to characterize the predictability of interest rates in a new form that highlights the connection between the evidence and stationary theories of bond pricing. Like others before us, we use the expectations hypothesis as a benchmark against which to measure the evidence.

The most common statement of the expectations hypothesis is that forward rates are expectations of future short rates:

$$f_t^n = E_t r_{t+n}, \tag{4}$$

a relation attributed to Hicks. Since this is easily contradicted by the data (average forward rates vary systematically by maturity), most people now understand the expectations hypothesis to include the possibility of a constant term premium p:

$$f_t^n = E_t r_{t+n} + p^n. \tag{4'}$$

The generic alternative is that term premiums vary through time:

$$f_t^n = E_t r_{t+n} + p_t^n. \tag{5}$$

Evidence against the expectations hypothesis, as defined by equation (4′), is therefore evidence that term premiums vary over time.

Equation (4) implies one of the most direct implications of the expectations hypothesis, that forward rates are martingales:

$$f_t^n = E_t f_{t+1}^{n-1}. \tag{6}$$

See Roll (1970, Chapter 4) or Sargent (1987, Section X.7). Given (6), or the weaker (4′), the regression

$$f_{t+1}^{n-1} - r_t = \text{constant} + c_n(f_t^n - r_t) + \text{residual} \tag{7}$$

has slope $c_n = 1$ for all maturities n. This "forward rate regression" is, as far as we know, new to the literature, but is a close relative of Fama (1984). The role of the term premium is apparent from the complementary regression,

$$f_{t+1}^{n-1} - f_t^n = \text{constant} + (c_n - 1)(f_t^n - r_t) + \text{residual},$$

a linear transformation of (7). When term premiums are constant, $E_t f_{t+1}^{n-1} - f_t^n = p^{n-1} - p^n$, a constant, and $c_n - 1 = 0$. Values of c_n different from one thus indicate that term premiums vary through time. Whether or not the expectations hypothesis holds, nonzero values of c_n in (7) indicate that forward rate spreads contain information that can be used to forecast future forward rates.

We report estimates of equation (7) in Table 1 and Figure 1 for all four sets of forward rate data. It should be no surprise that the data differ substantially from the expectations hypothesis. With the exception of the Unsmoothed Fama-Bliss data, estimated regression slopes are about one-half for $n = 1$, increase monotonically with maturity, and level off just below one.

Perhaps the most interesting feature of Figure 1 is that the largest deviations from the expectations hypothesis come at short maturities: the slopes are close to one for maturities of 24 months or longer. This feature of the data conforms well with theory: in a broad class of stationary bond-pricing models, the variance of term premiums approaches zero, and slopes of forward rate regressions approach one, as maturity increases. Related propositions are proved in different settings by Backus, Gregory, and Zin (1988, Proposition 2) and Dybvig, Ingersoll, and Ross (1996). Both are corollaries of a property that is well known in the Cox-Ingersoll-Ross model: that the variance of forward rates approaches zero with maturity. The evidence differs somewhat from this theoretical ideal: estimated slope coefficients are about 0.96 at long maturities, not the 1.00 suggested by the theory, with standard errors of 0.02 or smaller.

Our summary assessment of the expectations hypothesis — that the largest departures occur at short maturities — differs sharply in appearance from related work by Campbell and Shiller (1991) with "yield regressions," in which the largest departures are at long maturities. We suggest in Section 7 that the two approaches report similar information in different form. Note, though, that since yields are averages of forward rates, yield regressions involve term premiums over a range of maturities. For this reason, their slopes need not converge to one, even in theory.

Before turning to possible explanations, we consider two potential problems with the evidence. One is measurement error; see, for example, Bekaert, Hodrick, and Marshall (1996) and Stambaugh (1988). The effect of such error on estimated regression slopes depends on its form. Error in the short rate would likely push slope estimates toward one, but error in long forward rates would probably push them toward zero. To make this concrete, suppose our observations of forward rates $\hat{f}$ differ from "true" forward rates f in having measurement error η:

$$\hat{f}_t^i = f_t^i + \eta_t^i,$$

where η_t^i is independent of true forward rates, has variance σ_i^2, is uncorrelated with errors at different dates ($Corr(\eta_t^i, \eta_s^j) = 0$ for all dates $t \neq s$ and all maturities i,j), and has arbitrary correlation with contemporaneous errors ($Corr(\eta_t^i, \eta_t^j) = \rho_{ij}$). The population regression slope for true forward rates is

$$\begin{aligned} c_n &= \frac{Cov(f_{t+1}^{n-1} - f_t^0, f_t^n - f_t^0)}{Var(f_t^n - f_t^0)} \\ &= \frac{Cov(\hat{f}_{t+1}^{n-1} - \hat{f}_t^0, \hat{f}_t^n - \hat{f}_t^0) - [\sigma_0^2 - \rho_{0n}\sigma_0\sigma_n]}{Var(\hat{f}_t^n - \hat{f}_t^0) - [\sigma_0^2 + \sigma_n^2 - 2\rho_{0n}\sigma_0\sigma_n]}. \end{aligned} \tag{8}$$

Given estimates of error variances and correlations, we can estimate the impact on estimated regression slopes.

The difficulty is quantifying the error. One source of estimates is provided by McCulloch and Kwon (1993): standard deviations of each estimated forward rate. In Table 2, we report mean square standard deviations for a number of maturities for the period January 1970 to February 1991, the overlap of our samples. Standard deviations vary between 9 and 21 basis points, depending on the maturity.

Another approach is suggested by Bekaert, Hodrick, and Marshall (1996): standard deviations for differences between forward rates estimated by different methods, which are also reported in Table 2. Depending on its source, these numbers may either under- or over-estimate the magnitude of measurement error. If measurement error stems from the underlying data, and is therefore largely common across methods, we would expect the standard deviations computed this way to underestimate the measurement error. But if the error stems from the method, we might expect the standard deviations to overstate the error in one method alone. For the most part, the estimates point in the same direction as McCulloch and Kwon's: to measurement error with a standard deviation under 30 basis points for most maturities. The Unsmoothed Fama-Bliss estimates are a striking exception, with very large standard deviations for long maturities. We find this a likely explanation of the erratic behavior of regression estimates based on this data in Table 1.

We use (8) to estimate the impact of measurement error on estimated regression slopes. Consider the effect on the first and last regression slopes, c_1 and c_{120}. For the Smoothed Fama-Bliss data $Cov(\hat{f}_{t+1}^{n-1} - \hat{f}_t^0, \hat{f}_t^n - \hat{f}_t^0) = 0.0985$ for $n = 1$ and 3.196 for $n = 120$. Similarly, $Var(\hat{f}_t^n - \hat{f}_t^0) = 0.2162$ for $n = 1$ and 3.317 for $n = 120$. If we use the McCulloch and Kwon standard deviations from Table 2 and assume that errors are uncorrelated across maturities, then the error-adjusted slope estimates are

$$\begin{aligned} c_1 &= 0.462 \\ c_{120} &= 0.968. \end{aligned}$$

The impact is evidently small, particularly for long maturities. To get $c_{120} = 1$, for example, we need a standard deviation of 35 basis points for the error of f^{120}, a value more than double what we see in Table 2. Positive correlation between the errors typically reduces its overall effect. The effect of error is greater at the short end, largely because the relevant variances and covariances are smaller, but even here is not large enough to change one's interpretation of the evidence.

A second potential problem with the evidence is small sample bias. The methods used by Bekaert, Hodrick, and Marshall (1996) imply that small sample bias tends to push estimates of c_n above one, however, not below. Suppose, as they do, that the short rate is AR(1):

$$r_{t+1} = \text{constant} + \varphi r_t + \varepsilon_{t+1},$$

where $\{\varepsilon_{t+1}\}$ is a sequence of iid innovations in the short rate. The counterpart of their Proposition 1 is

$$E(\hat{c}_n) = 1 - \left(\frac{\varphi^{n-1}}{1 - \varphi^n}\right) [E(\hat{\varphi}) - \varphi],$$

where a "hat" again indicates an estimate of the underlying parameter. Since the small-sample bias in $\hat{\varphi}$ is approximately

$$E(\hat{\varphi}) - \varphi = -\frac{1 + 3\varphi}{T},$$

where T is the sample size, the bias in the regression slope under the expectations hypothesis is approximately

$$E(\hat{c}_n) - 1 = \left(\frac{\varphi^{n-1}}{1-\varphi^n}\right)\left(\frac{1+3\varphi}{T}\right). \tag{9}$$

Thus the estimated regression slope $\hat{c}_n$ is biased upward in small samples, in the opposite direction of observed departures from the expectations hypothesis. Small sample bias does not, therefore, appear to be the source of apparent predictable variation in term premiums.

We conclude — as did Bekaert, Hodrick, and Marshall (1996) in a related context — that both measurement error and small sample bias influence our estimated regression slopes, but that neither leads us to doubt the evidence that term premiums vary through time.

4 Affine Models

The next step is to consider models that might account for the time-varying term premiums we seem to see in the data. Multifactor Vasicek (1977) models can be ruled out immediately: they imply constant term premiums, and therefore cannot account for the apparent correlation of term premiums with forward rate spreads.

We consider, instead, a subset of Duffie and Kan's (1996) affine yield models in which conditional variances, and hence term premiums, vary over time. Fisher and Gilles (1996), Frachot and Lesne (1994), and Roberds and Whiteman (1996) study similar models for the same reason. Bond prices in these models are based on a vector of state variables z following

$$\begin{aligned} z_{t+1} - z_t &= (I-\Phi)(\theta - z_t) + \Sigma V(z_t)^{1/2}\varepsilon_{t+1} \\ &= K(\theta - z_t) + \Sigma V(z_t)^{1/2}\varepsilon_{t+1}, \end{aligned} \tag{10}$$

where $\{\varepsilon_t\} \sim \text{NID}(0, I)$, Φ is a stable matrix with positive diagonal elements, $K = I - \Phi$, Σ is a diagonal matrix with elements σ_i, and $V(z_t)$ is a diagonal matrix with elements $v_i(z_t) = z_{it}$. State prices are governed by a pricing kernel of the form

$$-\log m_{t+1} = \delta + \gamma^\top z_t + \lambda^\top V(z_t)^{1/2}\varepsilon_{t+1}. \tag{11}$$

These models are a subset of Duffie and Kan's affine-yield models in which we have restricted ourselves to nonconstant volatility functions v_i that depend only on the ith state variable. Translation into discrete time is now standard; see, for example, Campbell, Lo, and MacKinlay (1996, Chapter 11) and Sun (1992). The most common examples of this class are versions of the Cox-Ingersoll-Ross model with one or more factors: $\delta = 0$, Φ is diagonal, and $\gamma_i = 1 + \lambda_i^2/2$. The choice of γ_i is a normalization that defines the short rate as $r_t = \sum_i z_{it}$.

As with the Cox-Ingersoll-Ross model, we generally restrict the parameters to guarantee that the volatility functions v_i remain nonnegative. Define the set D of admissible states as those values of z for which volatility is nonnegative:

$$D = \{z : v_i(z) \geq 0 \text{ for all } i\}.$$

In our class of models, D consists of the positive orthant. Duffie and Kan (1996, Section 4) give sufficient conditions for z to remain in D. For each i and all $z \in D$ satisfying $v_i(z) = 0$ (the boundary of positive volatility), the "drift" [the conditional mean of (10)] must be sufficiently positive:

$$\kappa_{ii}\theta_i + \sum_{j \neq i} \kappa_{ij}(\theta_j - z_j) \geq \sigma_i^2/2 > 0,$$

where κ_{ij} is an element of K. This implies the usual Feller condition

$$\kappa_{ii}\theta_i \geq \sigma_i^2/2 \tag{12}$$

(set $z_j = \theta_j$). It also implies that off-diagonal elements of K are nonpositive (consider arbitrarily large values of z_j) and

$$\kappa_{ii}\theta_i + \sum_{j \neq i} \kappa_{ij}\theta_j \geq \sigma_i^2/2, \tag{13}$$

(set $z_j = 0$), which puts a bound on their magnitude. These statements are exact in continuous time, approximate in our discrete-time analog.

Conditions (12,13) are stronger than we need. Volatility functions and state variables remain positive under the weaker conditions

$$\kappa_{ii}\theta_i + \sum_{j \neq i} \kappa_{ij}(\theta_j - z_j) > 0,$$

for all i and $z \in D$, which implies $\kappa_{ii} = 1 - \varphi_{ii} > 0$ for all i and $\kappa_{ij} = -\varphi_{ij} < 0$ for all $i \neq j$. The weaker condition allows greater probability of values of the state variable and volatility functions near zero, but nevertheless requires the drift on the boundary of D to push z back into the interior. Both versions rule out unit roots in the state variables z: these models are stationary.

With this structure, bond prices are log-linear functions of the state variables z:

$$-\log b_t^n = A_n + B_n^\top z_t \tag{14}$$

for some parameters $\{A_n, B_n\}$. The parameters are derived using the pricing relation

$$b_t^{n+1} = E_t\left(m_{t+1} b_{t+1}^n\right), \tag{15}$$

starting with $b_t^0 = 1$. Application of (15) generates the recursions

$$\begin{aligned} A_{n+1} &= A_n + \delta + B_n^\top (I - \Phi)\theta \\ B_{i,n+1} &= \gamma_i + \sum_j B_{jn}\varphi_{ji} - (\lambda_i + B_{in}\sigma_i)^2/2, \end{aligned}$$

starting with $A_0 = 0$ and $B_0 = 0$.

We can now derive population values for the slopes of forward rate regressions. For this we need the unconditional variance of z, the solution to

$$\Gamma_0 = \Phi \Gamma_0 \Phi^\top + \Sigma V(\theta) \Sigma^\top,$$

where $\Sigma V(\theta)\Sigma^\top$ is a diagonal matrix with positive elements $\sigma_i^2 \theta_i$. The solution is

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

where $vec(A)$ is the vector formed from the columns of the matrix A. Forward rate regression slopes in this setting are then

$$c_n = \frac{[B_1 + B_n - B_{n+1}]^\top \Gamma_0 \,[B_1 - \Phi^\top (B_n - B_{n-1})]}{[B_1 + B_n - B_{n+1}]^\top \Gamma_0 \,[B_1 + B_n - B_{n+1}]}, \tag{16}$$

which may take on values other than one.

Although (16) is relatively opaque, its limiting behavior is not. Since B_n converges, the population value of the regression slope approaches one:

$$\lim_{n\to\infty} c_n = \frac{B_1^\top \Gamma_0 \, B_1}{B_1^\top \Gamma_0 \, B_1} = 1.$$

This is a generalization of the property we noted earlier for the Cox-Ingersoll-Ross model: The variance of forward rates falls to zero with maturity, so for long enough maturities we are effectively regressing $-r$ on itself.

5 One-Factor Models

In this section and the next, we examine the ability of one-factor and multifactor models, respectively, to account for the slopes of forward rate regressions. Although the empirical weaknesses of one-factor models have been clearly documented, they provide a useful source of intuition for more complex models. We show that while the one-factor Cox-Ingersoll-Ross model can reproduce the estimated slope of the first forward rate regression, it cannot simultaneously generate an upward sloping average yield or forward rate curve. However, a different one-factor model, which we term the "negative CIR," can account for both features of the data.

Consider, then, the one-factor Cox-Ingersoll-Ross model, a special case of equations (10,11) in which z is a scalar, $\delta = 0$, and $\gamma = 1 + \lambda^2/2$. Then $A_1 = 0$, $B_1 = 1$, $r_t = z_t$, and $B_2 = 1 + \varphi - \sigma(\lambda + \sigma/2)$. The population value of the first regression slope is

$$c_1 = \frac{1 - \varphi}{1 - \varphi + \sigma(\lambda + \sigma/2)}. \tag{17}$$

Depending on the parameter values, this can take on values greater than, equal to, or less than one.

We illustrate the consequences of different parameter choices with GMM estimates based on different moment conditions. In each case, the parameters are estimated by iterating on the weighting matrix (Campbell, Lo, and MacKinlay 1996, Appendix A.2, or Ogaki 1993). We compute the weighting matrix by the method of Newey and West with a maximum lag length of 12 (about which more will be said in the next section).

We first choose $(\theta, \sigma, \varphi, \lambda)$ to reproduce four moments: the mean, variance, and autocorrelation of the short rate, and the mean spread between the 10-year bond yield and the short rate. We report these and other moments in Table 3. Estimated parameter values are reported in Table 4 as Model A. These estimates provide a good approximation to the empirical literature on one-factor Cox-Ingersoll-Ross models: they approximate the properties of the short rate and the average upward slope of yield and forward rate curves. They do not, however, generate realistic slopes of forward rate regressions. The first one is $c_1 = 1.426$, which is the wrong side of one. The complete set of regression slopes for these parameter values is contrasted with the data in the upper panel of Figure 2 (solid line).

Alternatively, we could replace the fourth moment condition (the mean 10-year bond spread) with one based on the first forward rate regression slope (c_1). The parameter values implied by these moments are reported as Model B. The problem now is that the mean yield curve is decreasing (the dashed line in the lower panel of Figure 2).

This difficulty is a general one for the one-factor Cox-Ingersoll-Ross model. For the regression slope to lie between zero and one, we need $\sigma(\lambda + \sigma/2) > 0$. This implies, however, a downward-sloping mean forward rate curve. For example,

$$E(f^1 - f^0) = [1 - \sigma(\lambda + \sigma/2)]\, \theta - \theta = -\sigma(\lambda + \sigma/2)\, \theta,$$

which is negative when we choose parameters to reproduce the slope of the first forward rate regression. The parameters reported by Frachot and Lesne (1994) are similar: they reproduce various regression slopes but imply downward-sloping average yield and forward rate curves. The one-factor model cannot generate both an increasing mean forward rate curve (hence yield curve) and a regression slope between zero and one. Model C makes the same point statistically: when we use both moments (the regression slope and the mean 10-year bond spread) the J-statistic is 7.12, indicating substantial conflict between the model and these features of the data.

The source of this difficulty is the behavior of the term premium. Consider the behavior of the first forward rate spread, $f_t^1 - r_t$. The spread has two components, the expected change in the short rate and a term premium:

$$f_t^1 - r_t = E_t r_{t+1} - r_t + p_t^1;$$

see equation (5). If the two components move in the same direction, the implied regression slope is less than one, as we see in the data. But if we estimate the parameters to generate an upward-sloping average yield curve, as in Model A, the term premium varies inversely with the expected change in the short rate. The first term premium is

$$p_t^1 = \text{constant} + [\sigma(-\lambda + \sigma/2)]\, z_t.$$

The coefficient of z is positive when the mean forward rate curve is increasing. A decline in z raises the expected change in the short rate, but it reduces the term premium. As a result, the implied regression slope is not between zero and one. With our estimated parameter values it is greater than one (the solid line in Figure 2).

In a one-factor affine world, the regression slope requires a term premium that varies inversely with the short rate. A relatively simple way to accomplish this while retaining a positively sloped average yield curve is with what we term the "negative CIR model": $\gamma = -1 + \lambda^2/2$. Then $B_1 = -1$, $B_2 = -1 - \varphi - \sigma(\lambda - \sigma/2)$, and the short rate is $r_t = \delta - z_t$. The mean difference between the first two forward rates is

$$E(f^1 - f^0) = [1 + \sigma(-\lambda + \sigma/2)]\,\theta - \theta = [\sigma(-\lambda + \sigma/2)]\,\theta,$$

which is positive when the term in brackets is. The first term premium is now

$$p_t^1 = \text{constant} + [\sigma(-\lambda + \sigma/2)]\, z_t.$$

Since r rises when z falls, the term premium varies inversely with r when the term in brackets is positive. Under the same conditions, the first regression slope,

$$c_1 = \frac{1 - \varphi}{1 - \varphi + \sigma(-\lambda + \sigma/2)},$$

is between zero and one. Thus the model seems capable of resolving the tension between the regression slopes and the average slope of the forward rate curve.

This possibility is born out in Table 4, where estimated parameter values for the negative CIR model are reported as Model D. By design, the parameter values reproduce the mean, variance, and autocorrelation of the short rate, the mean 10-year bond spread, and the slope of the first forward rate regression.

In short, the one-factor Cox-Ingersoll-Ross model is incapable of accounting for both the evidence against the expectations hypothesis and the average upward slope of the yield curve. The negative Cox-Ingersoll-Ross model resolves this difficulty, but for a variety of reasons cannot be the last word on the subject. One reason we regard as relatively innocuous: since z takes on all positive values, the short rate is negative with positive probability. In our example the probability is small — roughly 2%. Like Duffie and Singleton (1996) in a similar context, we regard the possibility of negative interest rates a small price to pay for the convenience of a linear model. A more compelling reason is the one-factor structure: the evidence suggests that we need two factors, and possibly more, to explain the curvature of the mean yield curve (Gibbons and Ramaswamy 1993) and the dynamics of interest rates (Garbade 1986, Litterman and Scheinkman 1991, Stambaugh 1988). We therefore turn next to multifactor models.

6 Multifactor Models

We have seen that affine models can account for slopes of forward rate regressions, but that some models have difficulty reconciling these regressions with other properties of interest rates. The question becomes, then, not whether affine models can account for forward rate regressions, but rather what kinds of affine models provide the best approximation to the behavior of interest rates overall. Like our predecessors, we find that multifactor models provide better approximations than their one-factor counterparts. We find, in addition, that models with one negative factor approximate the data better than pure CIR models containing only positive factors.

We report estimates of five affine models in Table 5. Each is estimated by GMM using 11 moment conditions based, respectively, on the standard deviation and autocorrelation of the short rate; the mean and standard deviation of yield spreads for maturities 12, 60, and 120 months; and forward rate regression slopes for maturities 1, 6, and 12 months. We use the same weighting matrix for each model: the Newey-West covariance matrix implied by estimates of the three-factor Model I, which includes each of the other models as a special case and thus provides a common basis of comparison. The weighting matrix is approximately a fixed point for Model I: it both produces and is produced by the parameters reported in the table. The number of lags (12) is based on calculations suggested by Andrews (1991, equations 6.2 and 6.6): if residuals from the moment conditions are ARMA(1,1), the optimal "lag truncation parameter" for Model I is 11.

The primary difficulty in computing estimates of multifactor affine models is in identifying the separate mean parameters θ_i. We follow Chen and Scott (1993) in computing these parameters separately with a grid search. In all five models, we set the model's mean short rate to equal the sample mean. For Models F through I, we consider a grid of choices for θ's that satisfy this constraint. Given such a choice, we compute the other parameters by minimizing the GMM objective function (the J-statistic) in the usual way. We report estimates based on the θ's that produce the smallest function value. The objective function is extremely flat with respect to the θ's in the neighborhood of the estimates, suggesting that the data are not very informative about these parameters. Reported standard errors are conditional on the choice of θ's.

The five models in Table 5 indicate, we think, the benefits of a negative factor. Models E (CIR) and F (negative CIR) are the one-factor models from Table 4, estimated here with a larger set of moments. The J-statistics suggest that F approximates the data substantially better than E. If an estimated model generated the data, its J has approximately a chi-square distribution in large samples. Although in finite samples the distribution can be substantially different (Tauchen 1986), the J's for these two models highlight the weaknesses of the one-factor structure.

A negative factor is equally helpful in a two-factor setting. Model G is a two-factor CIR model like that studied by Roberds and Whiteman (1996). It fits substantially better than the one-factor CIR model (Model E), but worse than the one-factor negative CIR (Model F). Model H is a "mixed" model, with one positive and one negative factor: $\gamma_1 = -1+\lambda_1^2/2$ and $\gamma_2 = 1+\lambda_2^2/2$, making the short rate $r_t = z_{2t}-z_{1t}$. Its J-statistic is significantly smaller than G's. The clear implication of Models E to H is that a negative factor permits a better approximation to the data, on the whole, than traditional CIR models with comparable numbers of parameters. In this sense, the negative CIR and mixed models are useful additions to the literature.

Model I is a three-factor model with one negative and two positive factors: $\gamma_1 = -1 + \lambda_1^2/2$, $\gamma_2 = 1 + \lambda_2^2/2$, and $\gamma_3 = 1 + \lambda_3^2/2$, so the short rate is $r_t = z_{2t} + z_{3t} - z_{1t}$. Its primary role here is to provide a comprehensive "encompassing" model that we can use to assess the others. The large standard errors suggest that we are close to the limits of what this set of data and moment conditions can tell us. Its J-statistic implies, nevertheless, that it cannot account for all of the moment conditions.

Although Model H's J-statistic indicates that there remain significant tensions between model and data, we think it provides an informative interpretation of the evidence. Figure 3 compares the regression slopes implied by Model H to those we estimated directly and reported in Table 1. The properties of the data are represented by asterisks. The implications of the model are represented by three lines computed from 1000 random draws of the parameters. In each one, we draw parameter values from the asymptotic multivariate normal distribution for the parameters summarized by the standard errors in Table 5. The solid line is the median from these 1000 replications and the two dashed lines are the 5% and 95% quantiles. Since the quantiles are based on the asymptotic normal approximation, they likely understate the sampling variability in the model. The message is that the model provides a passable approximation to the estimated regression slopes at short maturities. At long maturities the numerical differences between model and data are small, but the sampling variability is even smaller.

The discrepancy between long-maturity regression slopes in the model and the data is a robust feature of these models when their parameters are chosen to reproduce observed properties of interest rates. Even when we add moment conditions for long-maturity regression slopes, estimated models imply regression slopes that are closer to one than we see in the data. We can reduce the rate of convergence of the slopes to one by choosing autoregressive parameters φ_{ii} closer to one, but this invariably raises the unconditional standard deviation of the short rate or yield spreads well beyond their sample values.

Model H also provides a new perspective on small sample bias. We estimate the bias in regression slopes by simulating the model. Using estimated parameter values, we generate 1000 samples of 312 observations each for the state variables z_1 and z_2, from which we calculate forward rates. For each sample, we use simulated forward rates to compute regression slopes. We estimate the small sample bias by the difference between the mean regression slope across the 1000 replications and the population regression slope given by equation (16). Figure 4 suggests that this bias (solid line) can be substantial, especially at short maturities. The methods proposed by Bekaert, Hodrick, and Marshall (1996), summarized in our equation (9), give a similar answer (dashed line). Both suggest that small sample bias can be an important problem for samples of the size used in this paper. Curiously, the positive bias makes the differences between theory and evidence even more striking.

We conclude that affine models with negative factors appear capable of reconciling slopes of forward rate regressions with other properties of interest rates. The J-statistics suggest that our best efforts leave some features of the data unexplained, but nevertheless provide reasonable approximations to the dynamics of interest rates and slopes of forward rate regressions at short maturities.

7 Other Regressions

We have focused our attention on regressions forecasting one-period changes in forward rates [equation (7)]. Here we consider other regressions found in the literature.

Yield Regressions

Some of the most popular assessments of interest rate dynamics involve "yield regressions" of the form

$$y_{t+1}^{n} - y_t^{n+1} = \text{constant} + d_n \left(\frac{y_t^{n+1} - r_t}{n} \right) + \text{residual}, \tag{18}$$

a relation that dates back at least to Roll (1970). Recent empirical studies include Bekaert, Hodrick, and Marshall (1996), Campbell and Shiller (1991), and Evans and

Lewis (1994) for the US and Bekaert, Hodrick, and Marshall (1995) and Hardouvelis (1994) for other countries. The expectations hypothesis [equation (4)] implies $d_n = 1$ for all maturities n, but the equation is otherwise quite different from the forward rate regression (7) studied earlier. Since yields are averages of forward rates [equation (3)], slopes for long maturities include information about short-maturity forward rates and term premiums.

In Table 6 and Figure 5, we report estimates of the slope of (18) for maturities between 1 and 120 months. Yield regressions look markedly different from forward rate regressions (compare Figures 1 and 5). While slopes approach one with forward rates, with yields the slopes get progressively more negative as we increase maturity. Note, too, that the Unsmoothed Fama-Bliss estimates are broadly similar to those based on other data sets. As we see in Panel B of Table 2, the averaging involved in computing yields brings the estimated measurement error in this method closer to the others than we saw with forward rates. The similarity of the estimates across data sets suggests, moreover, that measurement error is not the source of the substantial departures from the expectations hypothesis.

Comparison

Despite the obvious differences between yield and forward rate regressions, we think they capture similar information. This claim deserves comment, since the relation between them is not otherwise self-evident. For $n = 1$, the yield and forward rate regressions contain exactly the same information: $d_1 = 2c_1 - 1$. Thus a value of 0.4557 for c_1 corresponds to -0.0886 for d_1. These two numbers are equivalent ways of representing the same information. For other maturities, there is no exact correspondence. Consider $n = 2$. The yield regression can be rewritten as

$$\left[(f_{t+1}^0 - r_t) + (f_{t+1}^1 - r_t)\right] = \text{constant} + \left(\frac{d_2 + 2}{3}\right)\left[(f_t^1 - r_t) + (f_t^2 - r_t)\right] + \text{residual},$$

which contains the elements of forward rate regressions for $n = 1, 2$. However, d_2 cannot be computed from the first two forward rate slopes, c_1 and c_2, alone. We need to know, among other things, the covariance between $f_t^1 - r_t$ and $f_t^2 - r_t$. More simply: yield and forward rate regressions do not contain identical information for longer maturities.

To see the similarities between the regressions, we need to impose more structure on the problem. Suppose, for the sake of approximation, that forward rates are linear functions of a single state variable z:

$$f_t^n = \text{constant} + \alpha_n z_t, \tag{19}$$

starting with the normalization $\alpha_0 = 1$. Since our regressions involve future forward rates and yields, internal consistency requires, in addition, a linear relation for the conditional mean of z:

$$E_t z_{t+1} = \text{constant} + \varphi z_t. \tag{20}$$

One-factor affine models are a special case of (19,20) in which the uncountable parameter set $\{\varphi, \alpha_1, \alpha_2, \ldots\}$ depends on a finite number of primitive parameters. The most restrictive ingredient is probably the single factor. Although additional factors are clearly called for, the first factor typically accounts for at least 80% of the variance of yield changes (Garbade 1986, Litterman and Scheinkman 1991).

We view (19,20) as approximations that help us to clarify the relations between regressions. With them, we can derive regression slopes as functions of the α's and φ, and derive exact relations between regression slopes. With this structure, forward rate and yield regression slopes are

$$\begin{aligned} c_n &= \frac{\varphi\alpha_{n-1} - 1}{\alpha_n - 1} \\ d_n &= \frac{(n+1)\varphi A_{n-1} - nA_n}{A_n - (n+1)}, \end{aligned}$$

where $A_n = \sum_{i=0}^{n} \alpha_i$. The inverse relations are

$$\begin{aligned} \alpha_n &= \frac{\varphi\alpha_{n-1} - 1}{c_n} + 1 \\ \alpha_n &= \frac{A_{n-1}[(n+1)\varphi - n - d_n] + (n+1)d_n}{d_n + n}, \end{aligned}$$

each of which is easily computed.

Given equations (19,20), we can infer the slopes of yield regressions implied by forward rate regressions, and vice versa. In each case, we use slopes estimated with Smoothed Fama-Bliss data, set $\varphi = 0.959$ (the autocorrelation of the short rate in Table 3), and interpolate between missing maturities with a cubic spline. The forward rate slopes implied by yield regressions (the line in Figure 6) are very similar to those we estimate directly (the asterisks). Conversely, the yield regression slopes are similar to those implied by forward rate regressions (the solid line in Figure 7). Apparently forward rate and yield regressions contain approximately the same information.

One striking by-product of this exercise is a new perspective on the long end of the yield curve: that the large negative slopes of yield regressions for long maturities seem to correspond to the numerically small difference in forward rate slopes from one. Suppose, to the contrary, that forward rate slopes were one for maturities of 24 months or more. Then the implied slopes of yield regressions level off at about -1 (the dashed line in Figure 7). Apparently the increasingly negative slopes of yield regressions at long maturities are closely related to the small difference from one of slopes of forward rate regressions. Yield regressions simply report this information in a way that magnifies the numerical difference from the expectations hypothesis at long maturities.

Affine Interpretations

The next issue is how well affine models account for slopes of yield regressions. The affine bond models of Section 4 imply regression slopes of

$$d_n \quad = \quad \frac{[B_1 - (n+1)^{-1} B_{n+1}]^\top \Gamma_0 \, [\frac{n}{n+1} B_{n+1} - \Phi^\top B_n]}{[B_1 - (n+1)^{-1} B_{n+1}]^\top \Gamma_0 \, [B_1 - (n+1)^{-1} B_{n+1}]}$$

The limiting value is

$$\lim_{n \to \infty} d_n = \frac{B_1^\top \Gamma_0 \, (I - \Phi^\top) B}{B_1^\top \Gamma_0 \, B_1},$$

where B is the limit of B_n. Although the slope converges, it need not converge to one.

As with forward rate regression slopes, the two-factor Model H provides a good approximation at short maturities, but does less well at long maturities (see Figure 8). Both reflect, in our view, the same problem: the limiting features of this class of models clash with the evidence. Whether this represents a difficulty with the theory or a peculiar feature of a specific sample is hard to say.

Multiperiod Short Rate Regressions

Other popular approaches to interest rate dynamics are based on multiperiod forecasts of the short rate. One example is

$$r_{t+n} - r_t = \text{constant} + e_n (f_t^n - r_t) + \text{residual}, \tag{21}$$

which generates a direct test of (4). Estimates are reported by Fama (1984), Fama and Bliss (1987), and Mishkin (1988). Another example is

$$\sum_{i=1}^{n}\left(1-\frac{i}{n+1}\right)(r_{t+i}-r_{t+i-1}) = \text{constant} + g_n(y_t^{n+1}-r_t) + \text{residual}, \tag{22}$$

which has been estimated by Campbell and Shiller (1991) with US data and Bekaert, Hodrick, and Marshall (1995) with data for the US, the UK, and Germany. If we rewrite (22) as

$$(n+1)^{-1}\sum_{i=1}^{n}(r_{t+i}-r_t) = \text{constant} + g_n(n+1)^{-1}\sum_{i=1}^{n}(f_t^i-r_t) + \text{residual},$$

the similarity to (21) is apparent. The expectations hypothesis implies, in both cases, regression slopes of one.

We report estimates of equations (21) and (22) in Tables 7 and 8. Both are estimated over samples that leave room for future values used in constructing the dependent variable. Like our earlier forward rate regressions, equation (7), multiperiod short rate regressions exhibit pronounced differences from the expectations hypothesis at short maturities. They differ, however, in two respects: the coefficients do not increase monotonically with maturity and the standard errors are substantially larger at long maturities. The former is the result of cumulating term premiums over several maturities. With (say) $n = 3$, the regression incorporates the effects of term premiums for maturities 1, 2, and 3. The latter is the consequence, primarily, of the variability and overlap in multiperiod forecast errors: we do not have many independent observations of (say) 120-month forecasts of the short rate.

The increased sampling variability in these regressions means we cannot make any precise statements about the behavior of regression slopes at long maturities: the point estimates are numerically different from one, but the standard errors are 0.25 or larger. In contrast, the standard errors in our earlier regressions were in the neighborhood of 0.012. One advantage, then, of (7) over (21) is that we have more precise information about changes over one period than n. As Hodrick (1992) notes in a different context, it can be helpful for statistical reasons to divide multiperiod returns into one-period components.

8 Forecasting

The expectations hypothesis provides a clear (if counterfactual) rationale for using current forward rates to forecast future values. Under it, forward rates are informationally efficient forecasts of future forward rates of shorter maturities. Although the expectations hypothesis clashes with the data, we might nevertheless use estimated forward rate regressions (7) to forecast the future:

$$P_t f_{t+1}^{n-1} \quad = \quad r_t + \text{constant} + c_n (f_t^n - r_t),$$

where P_t denotes a forecast ("projection") made at date t. These forecasts can then be used to inform activist investment strategies. The estimates in Table 1 suggest that when the forward rate curve is steeper than average, forward rates will rise by less than implied by the expectations hypothesis, making investments in long bonds attractive. Conversely, when the forward rate curve is flat, investments in short bonds are attractive.

Without the expectations hypothesis, however, we have no reason to believe — even on theoretical grounds — that these forecasts are informationally efficient. Further, the multifactor models estimated in Section 6 suggest that two variables will forecast better than one. However, theory does not tell us what these variables are: in a two-factor model, any two linearly independent combinations of forward rates are as good as any others. One of the simplest extensions of (7) is

$$f_{t+1}^{n-1} - r_t = \text{constant} + c_{1n}(f_t^n - r_t) + c_{2n} r_t + \text{residual}. \tag{23}$$

Our logic was to include the same variable for all maturities and to use a level, rather than a spread, in the interest of diversification. Few of the details matter. We find, in this version and several others that we do not report, that the additional variable matters only at short maturities (see Table 9). Our interpretation mirrors Section 2: the expectations hypothesis remains a reasonable approximation at long maturities, but at short maturities there is some value in using an additional variable.

The effect of the extra variable is pictured in Figure 9 for the forward rate curve of December 1994. We see in the top panel that the forward rate curve at that time was humped: very steep at the short end then slightly decreasing. As a result, the expectations hypothesis implied increasing forward rates for short maturities. The regressions, on the other hand, imply more modest increases. The difference between the two for long maturities reflects the non-unit estimated slope coefficients for long maturities. Estimates of the multivariate regression, equation (23), are somewhat different. Since the short rate was below its mean, the negative estimates of c_{2n} lead to higher forecasts at very short maturities. The increase at long maturities reflects a combination of the negative estimated values of c_{2n} and the small reduction in c_{1n} relative to estimates of c_n in Table 1. We infer from this that regressions based on the expectations hypothesis may not be informationally efficient, and that we might do better with multivariate forecasts.

9 Final Remarks

We join a long list of contributors in reconsidering the widely-documented evidence of interest rate predictability. We summarize predictability in a new form that allows direct comparison with the limiting behavior of a large class of theoretical models. Like Frachot and Lesne (1994) and Roberds and Whiteman (1996), we construct affine models that approximate this feature of the data. Unlike them, we argue that models outside the Cox-Ingersoll-Ross class provide a better approximation to the overall behavior of interest rates. We argue, more generally, that predictable changes in yields and forward rates are an important source of information about interest rate dynamics that can be used to guide investment strategies and estimate parameters of theoretical models. Perhaps further work along the lines of Dai and Singleton (1996) will tell us how important these features of the data are relative to others, and which models account best for the many properties of interest rates.

The major outstanding issue is the economic interpretation of the interest rate behavior we document in the data and approximate with affine models. One interesting interpretation is provided by Zin (1997), who argues that dynamic responses to innovations in the two factors of a generic bond-pricing model are similar to the empirical responses of interest rates to monetary policy and the real economy, respectively. What is needed to complete this story is a connection between these responses and changes in the conditional variance of interest rates and the pricing kernel.

References

Andrews, Donald, 1991, "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica* 59, 817-858.

Backus, David, Allan Gregory, and Stanley Zin, 1988, "Risk premiums in the term structure: Evidence from artificial economies," *Journal of Monetary Economics* 24, 371-399.

Bekaert, Geert, Robert Hodrick, and David Marshall, 1995, "'Peso problem' explanations for term structure anomalies," manuscript, Stanford University, December.

Bekaert, Geert, Robert Hodrick, and David Marshall, 1996, "On biases in tests of the expectations hypothesis of the term structure of interest rates," *Journal of Financial Economics*, forthcoming.

Bliss, Robert, 1996, "Testing term structure estimation methods," *Advances in Futures and Options Research*, forthcoming.

Campbell, John, 1995, "Some lessons from the yield curve," *Journal of Economic Perspectives* 9 (Summer), 129-152.

Campbell, John, Andrew Lo, and Craig MacKinlay, 1996, *The Econometrics of Financial Markets*, Princeton NJ: Princeton University Press.

Campbell, John, and Robert Shiller, 1991, "Yield spreads and interest rate movements: A bird's eye view," *Review of Economic Studies* 58, 495-514.

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

Cox, John, Jonathan Ingersoll, and Stephen Ross, 1985, "A theory of the term structure of interest rates," *Econometrica* 53, 385-407.

Dai, Qiang, and Kenneth Singleton, 1997, "Specification analysis of affine term structure models," manuscript, Stanford University, February.

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

Duffie, Darrell, and Kenneth Singleton, 1995, "An econometric model of the term structure of interest rate swap yields," manuscript, Stanford University, September; forthcoming, *Journal of Finance*.

Dybvig, Philip, Jonathan Ingersoll, and Stephen Ross, 1996, "Long forward and zero-coupon rates can never fall," *Journal of Business* 69, 1-25; early versions of this paper were in circulation by 1987.

El Kouri, Nicole, Antoine Frachot, and Helyette Geman, 1996, "A note on the behavior of long zero coupon rates in a no arbitrage framework," manuscript, Paris, May.

Elton, E.J., and T.C. Green, 1996, "Tax and liquidity effects in pricing government bonds," manuscript, Stern School of Business, New York University.

Evans, Martin, and Karen Lewis, 1994, "Do stationary risk premia explain it all? Evidence from the term structure," *Journal of Monetary Economics* 33, 285-318.

Fama, Eugene, 1984, "The information in the term structure," *Journal of Financial Economics* 13, 509-528.

Fama, Eugene, and Robert Bliss, 1987, "The information in long-maturity forward rates," *American Economic Review* 77, 680-692.

Fisher, Mark, and Christian Gilles, 1996, "Term premia in an exponential-affine economy," manuscript, Board of Governors.

Frachot, Antoine, and Jean-Philippe Lesne, 1994, "Expectations hypothesis and stochastic volatilities," manuscript, Banque de France, March.

Garbade, Kenneth, 1986, "Modes of fluctuation in bond yields," Topics in Money and Securities Markets No. 20, Bankers Trust, June; reprinted in K. Garbade, ed., *Fixed Income Analytics*, Cambridge: MIT Press, 1996.

Gibbons, Michael, and Krishna Ramaswamy, 1993, "A test of the Cox, Ingersoll, Ross model of the term structure," *Review of Financial Studies* 6, 619-658.

Hardouvelis, Gikas, 1994, "The term structure spread and future changes in long and short rates in the G7 countries," *Journal of Monetary Economics* 33, 255-283.

Hodrick, Robert, 1992, "Dividend yields and expected stock returns: Alternative procedures for inference and measurement," *Review of Financial Studies* 5, 357-386.

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

McCulloch, J. Huston, and Heon-Chul Kwon, 1993, "US term structure data, 1947-91," manuscript and computer diskettes, Ohio State University.

Mishkin, Frederick, 1988, "The information in the term structure: Some further results," *Journal of Applied Econometrics* 3, 307-314.

Ogaki, Masao, 1993, "Generalized method of moments: Econometric applications," in G.S. Maddala, C.R. Rao, and H.D. Vinod, eds., *Handbook of Statistics, Volume 11*, New York: North-Holland, 455-488.

Roberds, William, and Charles Whiteman, 1996, "Endogenous term premia and anomalies in the term structure of interest rates: Explaining the predictability smile," manuscript, Federal Reserve Bank of Atlanta, October.

Roll, Richard, 1970, *The Behavior of Interest Rates*, New York: Basic Books.

Sargent, Thomas, 1987, *Macroeconomic Theory (Second Edition)*, Orlando: Academic Press.

Stambaugh, Robert, 1988, "The information in forward rates: Implications for models of the term structure," *Journal of Financial Economics* 21, 41-70.

Sun, Tong-sheng, 1992, "Real and nominal interest rates: A discrete-time model and its continuous-time limit," *Review of Financial Studies* 5, 581-611.

Tauchen, George, 1986, "Statistical properties of GMM estimators of structural parameters obtained from financial market data," *Journal of Business and Economic Statistics* 4, 397-416.

Vasicek, Oldrich, 1977, "An equilibrium characterization of the term structure," *Journal of Financial Economics* 5, 177-188.

Zin, Stanley, 1997, "Discussion of Evans and Marshall," Carnegie-Rochester conference on Public Policy, November.

Table 1
Forward Rate Regressions in Four Data Sets

Maturity n	Smoothed Fama-Bliss	Unsmoothed Fama-Bliss	McCulloch Cubic Spline	Extended Nelson-Siegel
1	0.4557	0.5241	0.4984	0.5184
	(0.1851)	(0.1063)	(0.1760)	(0.1054)
3	0.7308	0.5758	0.7176	0.7979
	(0.0916)	(0.0872)	(0.0875)	(0.0932)
6	0.7971	0.6211	0.7833	0.8280
	(0.0570)	(0.0626)	(0.0569)	(0.0579)
9	0.8513	0.7594	0.8540	0.8703
	(0.0457)	(0.0524)	(0.0461)	(0.0461)
12	0.8913	0.7643	0.8916	0.9036
	(0.0393)	(0.0567)	(0.0373)	(0.0398)
24	0.9456	0.5843	0.9357	0.9500
	(0.0239)	(0.0591)	(0.0270)	(0.0251)
36	0.9576	0.6758	0.9544	0.9587
	(0.0172)	(0.0609)	(0.0167)	(0.0178)
48	0.9618	0.7630	0.9452	0.9612
	(0.0141)	(0.0313)	(0.0183)	(0.0143)
60	0.9635	0.5765	0.9565	0.9624
	(0.0124)	(0.0668)	(0.0122)	(0.0125)
84	0.9644	0.7086	0.9677	0.9636
	(0.0109)	(0.0639)	(0.0090)	(0.0108)
120	0.9634	0.6910	0.9655	0.9643
	(0.0102)	(0.0697)	(0.0102)	(0.0101)

Entries are estimated slope parameters c_n from

$$f_{t+1}^{n-1} - r_t = \text{constant} + c_n(f_t^n - r_t) + \text{residual},$$

where f^n is the n-month-ahead continuously-compounded one-month forward rate and $r_t = f_t^0$ is the short rate, both expressed as annual percentages. Forward rates were computed with data and programs supplied by Robert Bliss and come in four versions: Smoothed Fama-Bliss, Unsmoothed Fama-Bliss, McCulloch Cubic Spline, and Extended Nelson-Siegel. The data are monthly. For most entries, dates t run from January 1970 to November 1995 (311 observations). The exceptions concern the McCulloch Cubic Spline data, for which the starting dates are August 1971 for $n = 84$ and November 1971 for $n = 120$. Numbers in parentheses are Newey-West standard errors (6 lags).

Table 2
Estimated Standard Deviations of Measurement Error

Maturity n	McCulloch and Kwon	Difference from Smoothed-Fama Bliss: Unsmoothed Fama-Bliss	McCulloch Cubic Spline	Extended Nelson-Siegel
		A. Forward Rates		
0	0.11	0.16	0.22	0.19
1	0.13	0.23	0.19	0.15
3	0.17	0.30	0.12	0.09
6	0.18	0.46	0.18	0.07
9	0.15	0.45	0.16	0.08
12	0.21	1.13	0.18	0.08
24	0.16	1.71	0.24	0.07
36	0.13	1.62	0.21	0.06
48	0.13	2.03	0.26	0.07
60	0.09	2.01	0.24	0.08
84	0.10	2.91	0.30	0.12
120	0.13	3.52	0.65	0.18
		B. Yields		
1	0.11	0.16	0.22	0.19
3	0.05	0.07	0.10	0.03
6	0.04	0.06	0.06	0.04
9	0.05	0.08	0.07	0.03
12	0.04	0.09	0.07	0.03
24	0.03	0.08	0.06	0.04
36	0.02	0.07	0.06	0.04
48	0.02	0.09	0.07	0.03
60	0.02	0.08	0.08	0.03
84	0.02	0.12	0.09	0.04
120	0.03	0.17	0.12	0.06

Entries are estimated standard deviations of measurement error in forward rates and yields of different maturities, measured as annual percentages (the number 0.11, for example, corresponds to 11 basis points). The column labelled McCulloch and Kwon is the root mean square of the estimated standard deviations reported by McCulloch and Kwon (1993) over the period January 1970 to February 1992. The remaining three columns are standard deviations of differences in forward rates from Smoothed Fama-Bliss estimates. The sample period in most cases in January 1970 to December 1995. The exceptions concern the McCulloch Cubic Spline data, which start in August 1971 for $n = 84$ and November 1971 for $n = 120$.

Table 3
Properties of Forward Rates and Yields

Maturity	Mean	Std Deviation	Autocorrelation
	A. Forward Rates		
0	6.683	2.703	0.959
1	7.098	2.822	0.969
3	7.469	2.828	0.968
6	7.685	2.701	0.966
9	7.812	2.487	0.966
12	7.921	2.495	0.969
24	8.274	2.264	0.977
36	8.498	2.135	0.979
48	8.632	2.059	0.980
60	8.714	2.013	0.980
84	8.802	1.967	0.980
120	8.858	1.946	0.980
	B. Yields		
1	6.683	2.703	0.959
3	7.039	2.781	0.971
6	7.297	2.774	0.971
9	7.441	2.725	0.970
12	7.544	2.672	0.970
24	7.819	2.495	0.973
36	8.009	2.373	0.976
48	8.148	2.287	0.977
60	8.253	2.224	0.978
84	8.398	2.141	0.979
120	8.529	2.073	0.981

Entries are sample moments of continuously-compounded forward rates and yields constructed with data and programs supplied by Robert Bliss (Smoothed Fama-Bliss method). The data are monthly, January 1970 to December 1995 (312 observations). Mean is the sample mean, St Deviation the sample standard deviation, and Autocorrelation the first autocorrelation.

Table 4
Estimates of One-Factor Affine Models

Parameter	A	B	C	D
γ	$1+\lambda^2/2$	$1+\lambda^2/2$	$1+\lambda^2/2$	$-1+\lambda^2/2$
δ				0.00866 (0.00120)
θ	0.00557 (0.00043)	0.00557 (0.00043)	0.00543 (0.00043)	0.00309 (0.00114)
σ	0.00854 (0.00198)	0.00853 (0.00198)	0.00641 (0.00188)	0.01149 (0.00413)
$\lambda\sigma$	-0.0123 (0.00398)	0.0489 (0.0493)	-0.0111 (0.00379)	0.0509 (0.0511)
φ	0.9590 (0.0180)	0.9590 (0.0180)	0.961 (0.0182)	0.9586 (0.0180)
J-statistic	0.00	0.00	9.69	0.00
Deg of Fr	0	0	1	0
p	—	—	0.002	—
c_1	1.426	0.448	1.401	0.449
c_{12}	1.028	0.974	1.027	0.975
$E(y^{120}-y^1)$	1.846	-3.366	1.689	1.849
$Auto(r)$	0.959	0.959	0.961	0.9586

Parameter values were estimated by GMM, as described in the text, using yields and forward rates estimated by the Smoothed Fama-Bliss method. Numbers in parentheses are standard errors. J is the Hansen's J-statistic, Deg of Fr is its degrees of freedom, and p its marginal significance level. The other statistics are regression slopes for n equal to 1 and 12, the mean 10-year yield spread, and the first autocorrelation of the short rate. The sample period runs from January 1970 to November 1995 (311 observations). Weighting matrices were computed by the Newey-West method (12 lags). The moment conditions vary across models. Model A (benchmark CIR): the mean, standard deviation, and autocorrelation of the short rate, and the mean of the 10-year bond yield; Model B (regression-based CIR): the mean, standard deviation, and autocorrelation of the short rate, and the slope of the first forward rate regression; Model C (composite CIR): the union of the Models A and B; and Model D (negative CIR): same as C.

Table 5
Estimates of Multifactor Affine Models

Parameter	E	F	G	H	I
γ_1	$1+\lambda_1^2/2$	$-1+\lambda_1^2/2$	$1+\lambda_1^2/2$	$-1+\lambda_1^2/2$	$-1+\lambda_1^2/2$
δ		0.00857			
θ_1	0.00557	0.00300	0.00307	0.00150	0.00155
θ_2			0.00250	0.00707	0.00442
θ_3					0.00270
σ_1	0.00184	0.00274	0.00006	0.00190	0.00212
	(0.00014)	(0.00010)	(0.00534)	(0.00014)	(0.00026)
σ_2			0.00163	0.00140	0.00154
			(0.00028)	(0.00014)	(0.00151)
σ_3					0.00008
					(0.0340)
$\lambda_1\sigma_1$	–0.0118	0.00376	–0.0613	0.0748	0.0381
	(0.0021)	(0.0020)	(0.0168)	(0.0180)	(0.0093)
$\lambda_2\sigma_2$			0.0328	–0.0125	–0.0310
			(0.0110)	(0.0055)	(0.165)
$\lambda_3\sigma_3$					–0.000009
					(0.569)
φ_{11}	0.914	0.994	0.813	0.996	0.998
	(0.0084)	(0.0016)	(0.0519)	(0.0019)	(0.0017)
φ_{22}			0.916	0.653	0.758
			(0.0221)	(0.0811	(0.735)
φ_{33}					0.145
					(218.1)
J-statistic	221.25	53.80	125.98	17.54	7.58
Deg of Fr	8	7	5	5	2
p	$< 10^{-16}$	10^{-9}	10^{-16}	0.0036	0.0226
c_1	1.158	0.144	0.720	0.437	0.382
c_{12}	1.008	0.945	0.989	0.956	0.943
$E(y^{120}-y^1)$	0.937	2.498	0.892	1.838	2.149
$Auto(r)$	0.914	0.994	0.916	0.984	0.994

Parameter values were estimated by GMM, as described in the text, using yields and forward rates estimated by the Smoothed Fama-Bliss method. Numbers in parentheses are standard errors. J is the Hansen's J-statistic, Deg of Fr is its degrees of freedom, and p its marginal significance level. The other statistics are regression slopes for n equal to 1 and 12, the mean 10-year yield spread, and the first autocorrelation of the short rate. The sample period runs from January 1970 to November 1995 (311 observations). Models differ in number of factors and choice of γ_1 ($\gamma_i = 1 + \lambda_i^2/2$ for $i = 2, 3$). The same moment conditions and weighting matrix were used to estimate all five models. The 11 moment conditions are based on: the standard deviation and autocorrelation of the short rate, the mean and standard deviation of the spreads between long yields and the short rate ($y^n - r$ for $n = 12, 60, 120$) and the slopes of forward regressions (c_n for $n = 1, 6, 12$). The weighting matrix is based on Model I and was computed by the Newey-West method (12 lags). All reported statistics are conditional on the θ's, which were chosen by grid search to minimize the J-statistic subject to the restriction that the model's mean short rate equals the sample mean.

Table 6
Yield Regressions in Four Data Sets

Maturity n	Smoothed Fama-Bliss	Unsmoothed Fama-Bliss	McCulloch Cubic Spline	Extended Nelson-Siegel
1	−0.0886	0.0488	−0.0031	0.0369
	(0.3702)	(0.2127)	(0.3519)	(0.2107)
3	−0.4284	−0.6325	−0.4311	−0.1521
	(0.4808)	(0.6162)	(0.4751)	(0.3930)
6	−0.8828	−0.9305	−0.8305	−0.4979
	(0.6398)	(0.6911)	(0.6324)	(0.6252)
9	−1.2280	−1.3677	−1.0717	−0.8000
	(0.7380)	(0.7801)	(0.7512)	(0.7499)
12	−1.4248	−1.4753	−1.1255	−0.9901
	(0.8249)	(0.8109)	(0.8513)	(0.8440)
24	−1.7048	−1.9225	−1.4783	−1.3158
	(1.1202)	(1.1392)	(1.1100)	(1.1545)
36	−1.9100	−1.8528	−1.7818	−1.5482
	(1.2954)	(1.3005)	(1.3176)	(1.3508)
48	−2.1469	−2.6088	−2.1295	−1.8134
	(1.4180)	(1.3997)	(1.4420)	(1.4816)
60	−2.4333	−2.6205	−2.5064	−2.1344
	(1.5190)	(1.5304)	(1.5549)	(1.5834)
84	−3.0959	−3.4028	−2.9845	−2.8455
	(1.7047)	(1.6684)	(1.6534)	(1.7594)
120	−4.1729	−4.3454	−3.7285	−3.9662
	(1.9847)	(1.9217)	(2.0116)	(2.0196)

Entries are estimated slope parameters d_n from

$$y_{t+1}^{n} - y_t^{n+1} = \text{constant} + d_n \left(\frac{y_t^{n+1} - r_t}{n} \right) + \text{residual},$$

where y^n is the n-month continuously-compounded yield and $r_t = y_t^1$ is the short rate, both expressed as annual percentages. Yields were computed with data and programs supplied by Robert Bliss and come in four versions: Smoothed Fama-Bliss, Unsmoothed Fama-Bliss, McCulloch Cubic Spline, and Extended Nelson-Siegel. The data are monthly. For most entries, dates t run from January 1970 to November 1995 (311 observations). The exceptions concern the McCulloch Cubic Spline data, for which the starting dates are August 1971 for n = 84 and November 1971 for n = 120. Numbers in parentheses are Newey-West standard errors (6 lags).

Table 7
Multiperiod Regressions 1 in Four Data Sets

Maturity n	Smoothed Fama-Bliss	Unsmoothed Fama-Bliss	McCulloch Cubic Spline	Extended Nelson-Siegel
1	0.4557	0.5241	0.4984	0.5184
	(0.1772)	(0.1431)	(0.1638)	(0.1062)
3	0.4108	0.2682	0.3790	0.5814
	(0.1777)	(0.1944)	(0.1884)	(0.1728)
6	0.3624	0.3877	0.4199	0.4833
	(0.1835)	(0.1474)	(0.2014)	(0.1969)
9	0.4100	0.4598	0.4213	0.4788
	(0.1863)	(0.1695)	(0.1915)	(0.1857)
12	0.4590	0.1649	0.4139	0.5310
	(0.2092)	(0.1248)	(0.2049)	(0.2015)
24	0.6097	0.2788	0.6496	0.6319
	(0.3999)	(0.1923)	(0.3883)	(0.3981)
36	0.8431	0.3545	0.7409	0.8129
	(0.2942)	(0.1623)	(0.2867)	(0.2915)
48	0.9773	0.3070	1.0079	0.9045
	(0.2785)	(0.1385)	(0.2901)	(0.2656)
60	0.8513	0.5097	0.7731	0.7784
	(0.2825)	(0.1754)	(0.2618)	(0.2731)
84	0.6261	0.0785	0.6023	0.5535
	(0.5050)	(0.1888)	(0.4679)	(0.4922)
120	0.5988	0.5028	0.7324	0.5492
	(0.2538)	(0.1038)	(0.1286)	(0.2532)

Entries are estimated slope parameters e_n from

$$r_{t+n} - r_t = \text{constant} + e_n\,(f_t^n - r_t) + \text{residual},$$

where f^n is the n-month continuously-compounded forward rate and $r_t = f_t^0$ is the short rate, both expressed as annual percentages. Forward rates were computed with data and programs supplied by Robert Bliss and come in four versions: Smoothed Fama-Bliss, Unsmoothed Fama-Bliss, McCulloch Cubic Spline, and Extended Nelson-Siegel. The data are monthly. For most entries, dates t run from January 1970 to n months prior to December 1995 ($312 - n$ observations). The exceptions concern the McCulloch Cubic Spline data, for which the starting dates are August 1971 for $n = 84$ and November 1971 for $n = 120$. Numbers in parentheses are Newey-West standard errors (n lags).

Table 8
Multiperiod Regressions 2 in Four Data Sets

Maturity n	Smoothed Fama-Bliss	Unsmoothed Fama-Bliss	McCulloch Cubic Spline	Extended Nelson-Siegel
1	0.4557	0.5241	0.4984	0.5184
	(0.1772)	(0.1431)	(0.1638)	(0.1062)
3	0.4099	0.3625	0.4042	0.5602
	(0.1833)	(0.2283)	(0.1913)	(0.1441)
6	0.3848	0.4063	0.3831	0.5595
	(0.1529)	(0.1764)	(0.1683)	(0.1670)
9	0.3918	0.4184	0.4229	0.5384
	(0.1704)	(0.1904)	(0.1899)	(0.1876)
12	0.4280	0.4216	0.4621	0.5563
	(0.1763)	(0.1826)	(0.1950)	(0.1900)
24	0.5393	0.5423	0.5314	0.6230
	(0.2437)	(0.2562)	(0.2360)	(0.2347)
36	0.6102	0.5827	0.5733	0.6536
	(0.2635)	(0.2686)	(0.2581)	(0.2589)
48	0.6989	0.6433	0.6538	0.7123
	(0.2475)	(0.2496)	(0.2520)	(0.2415)
60	0.7622	0.7352	0.7255	0.7541
	(0.2352)	(0.2346)	(0.2411)	(0.2253)
84	0.6824	0.6741	0.6199	0.6592
	(0.2524)	(0.2712)	(0.2356)	(0.2459)
120	0.7191	0.7665	0.7086	0.6782
	(0.2582)	(0.2566)	(0.2274)	(0.2519)

Entries are estimated slope parameters g_n from

$$\sum_{i=1}^{n}\left(1-\frac{i}{n+1}\right)(r_{t+i}-r_{t+i-1}) = \text{constant} + g_n\left(y_t^{n+1}-r_t\right) + \text{residual},$$

where y^n is the n-month continuously-compounded yield and $r_t = y_t^1$ is the short rate, both expressed as annual percentages. Forward rates were computed with data and programs supplied by Robert Bliss and come in four versions: Smoothed Fama-Bliss, Unsmoothed Fama-Bliss, McCulloch Cubic Spline, and Extended Nelson-Siegel. The data are monthly. For most entries, dates t run from January 1970 to n months prior to December 1995 ($312-n$ observations). The exceptions concern the McCulloch Cubic Spline data, for which the starting dates are August 1971 for $n = 84$ and November 1971 for $n = 120$. Numbers in parentheses are Newey-West standard errors (n lags).

Table 9
Multivariate Forward Rate Regressions

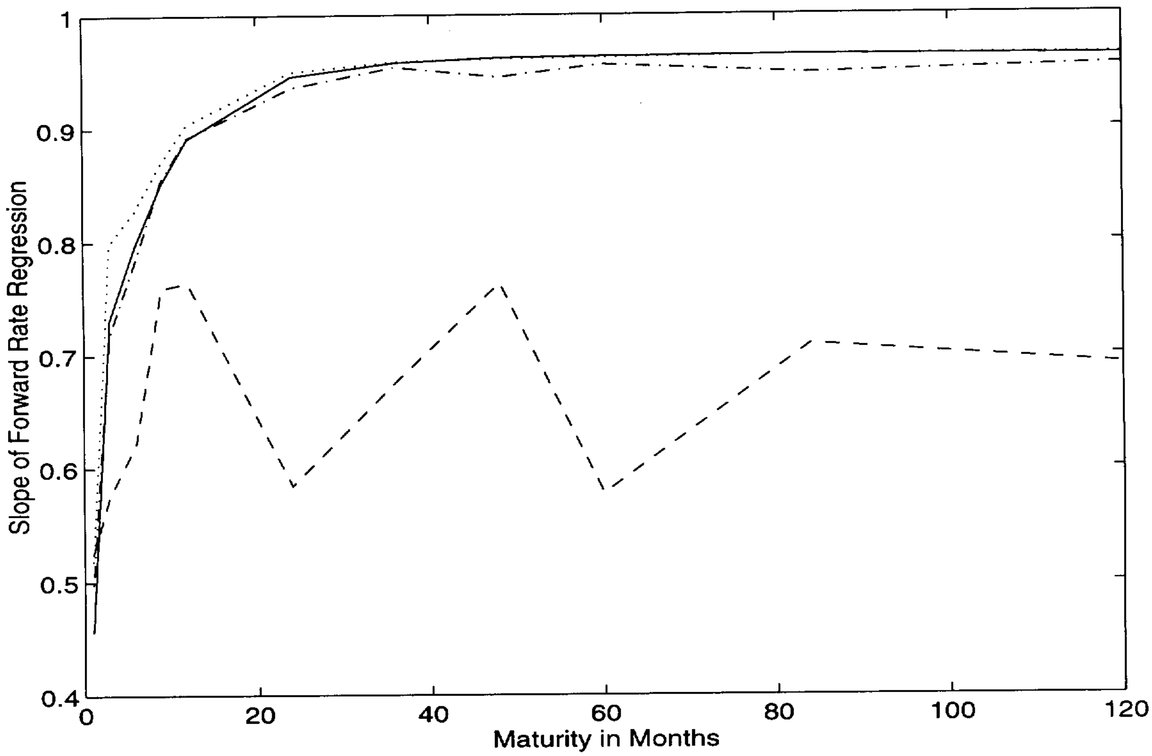
Maturity n	Constant	c_{1n}	c_{2n}	R^2
1	0.1451	0.5118	−0.0549	0.113
	(0.1412)	(0.1801)	(0.0215)	
3	0.0919	0.7322	−0.0048	0.326
	(0.1475)	(0.0882)	(0.0211)	
6	0.1951	0.7932	−0.0076	0.477
	(0.1382)	(0.0581)	(0.0179)	
9	0.2342	0.8389	−0.0150	0.614
	(0.1305)	(0.0462)	(0.0166)	
12	0.2213	0.8755	−0.0166	0.722
	(0.1235)	(0.0400)	(0.0156)	
24	0.1591	0.9324	−0.0126	0.897
	(0.1050)	(0.0265)	(0.0123)	
36	0.1528	0.9455	−0.0115	0.937
	(0.1005)	(0.0214)	(0.014)	
48	0.1643	0.9491	−0.0122	0.950
	(0.0998)	(0.0192)	(0.0112)	
60	0.1777	0.9498	−0.0131	0.956
	(0.1000)	(0.0182)	(0.0112)	
84	0.1954	0.9494	−0.0144	0.961
	(0.0999)	(0.0174)	(0.0112)	
120	0.2104	0.9476	−0.0153	0.962
	(0.0961)	(0.0165)	(0.0107)	

Entries are estimated parameters from the multivariate forward rate regression,

$$f_{t+1}^{n-1} - r_t = \text{constant} + c_{1n}(f_t^n - r_t) + c_{2n} r_t + \text{residual},$$

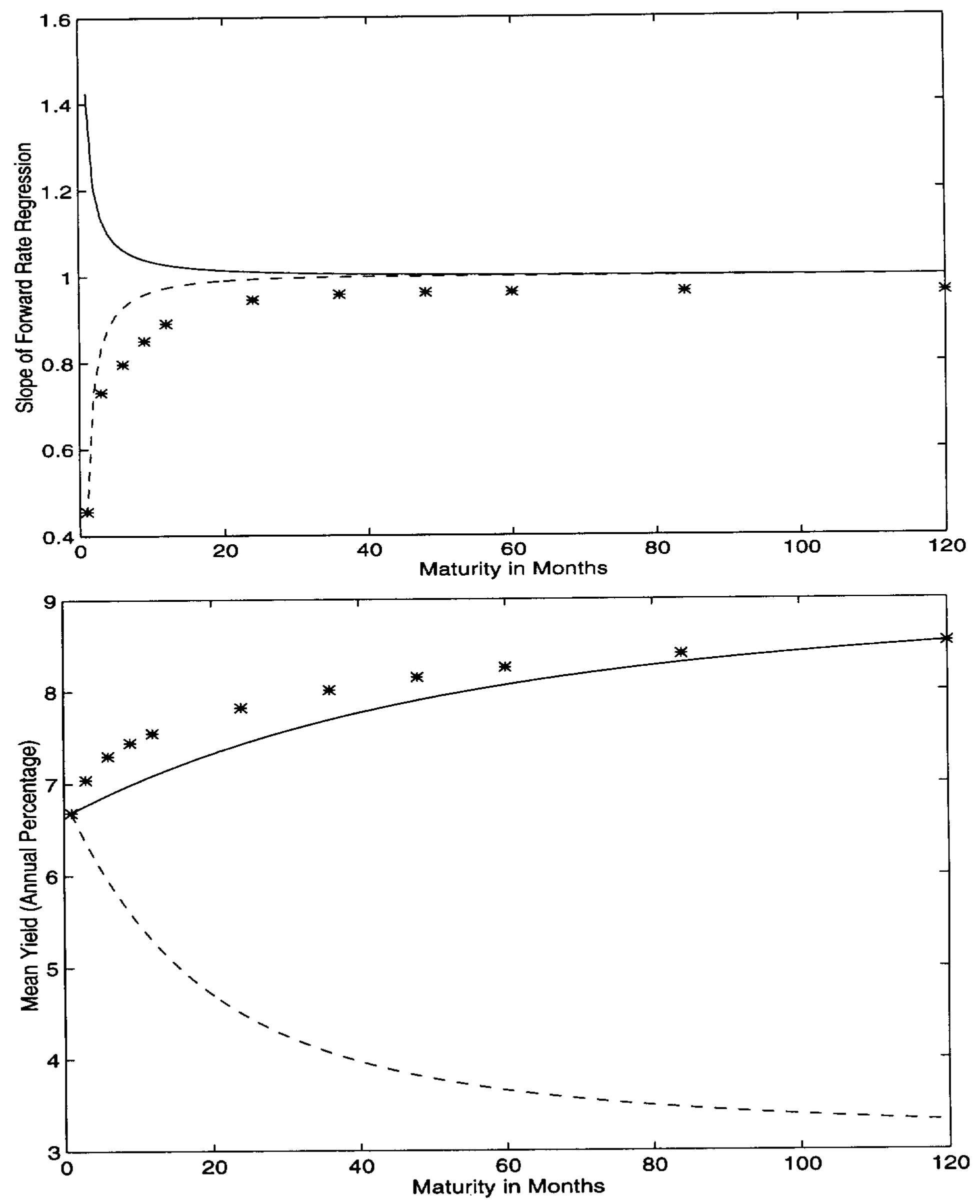
based on Smoothed Fama-Bliss estimates of forward rates. Definitions are provided in Table 1. Dates t run from January 1970 to November 1995 (311 observations). Numbers in parentheses are Newey-West standard errors (6 lags).

Figure 1
Slopes of Forward Rate Regressions in Four Data Sets



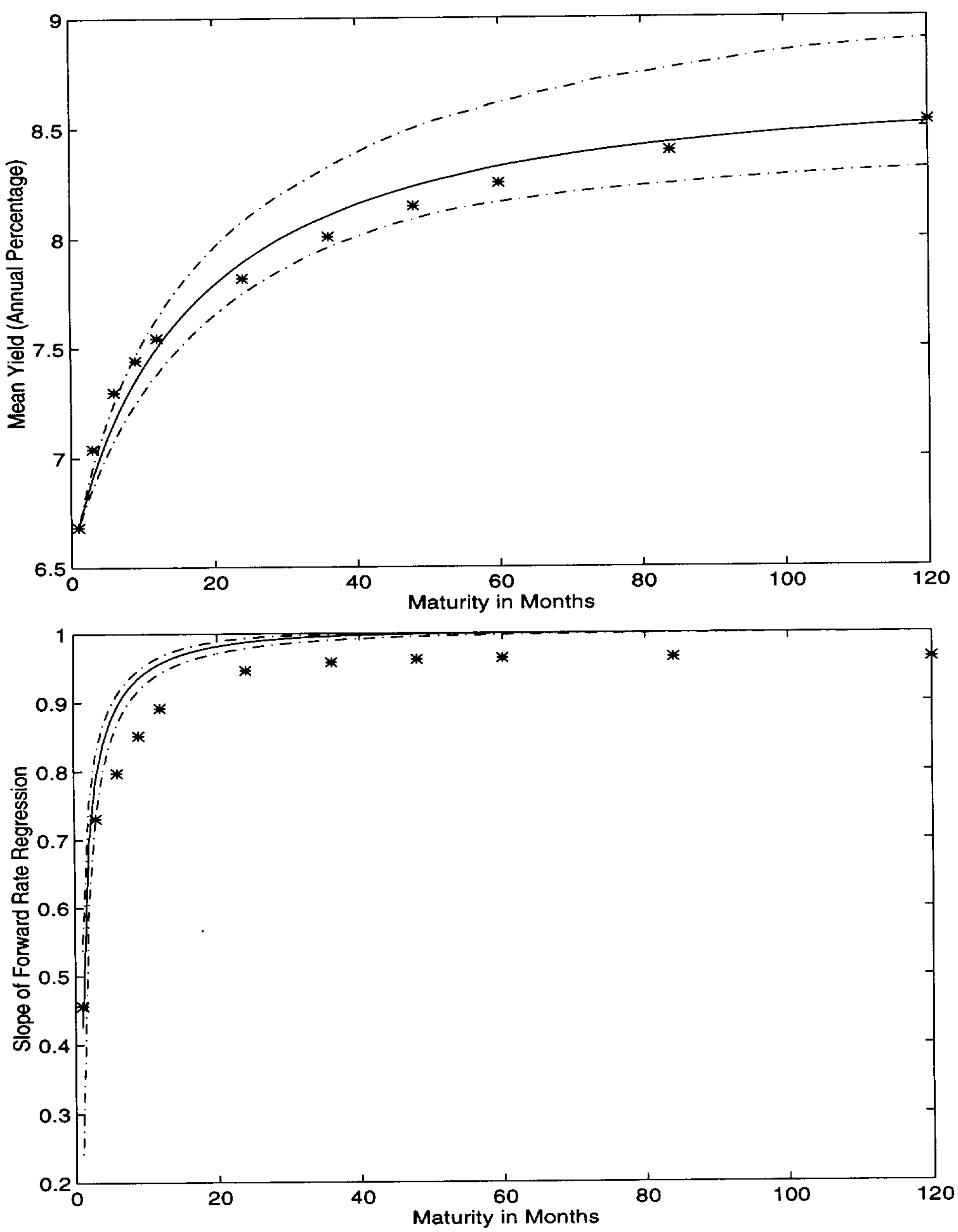
Lines represent the forward rate regression slopes c_n reported in Table 1 for forward rates estimated by four different methods. The solid line is based on data estimated by the Smoothed Fama-Bliss method, the dashed line (the markedly different one) the Unsmoothed Fama-Bliss method, the dash-dotted line the McCulloch Cubic Spline method, and the dotted line the Extended Nelson-Siegel method.

Figure 2
Properties of the One-Factor Cox-Ingersoll-Ross Model



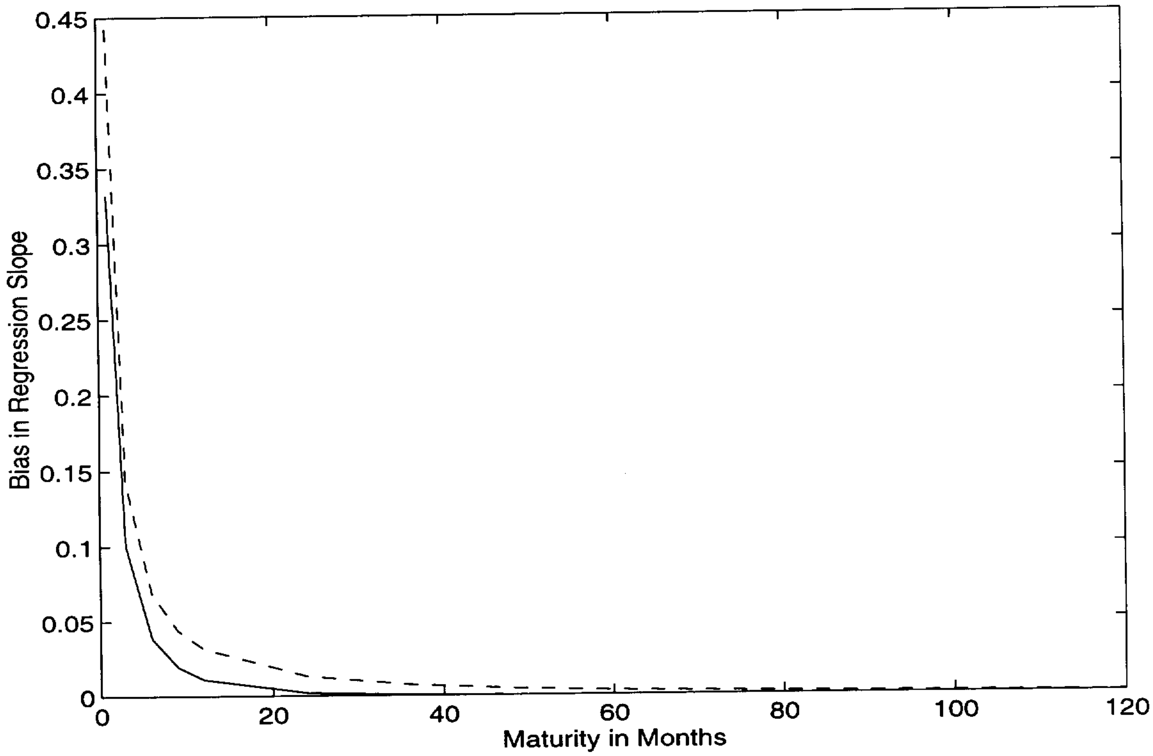
The two panels report forward rate regression slopes and mean yields in the data (the asterisks) and in two versions of the one-factor Cox-Ingersoll-Ross model, Models A and B from Table 4. Model A (the solid line) is estimated to reproduce the average 10-year bond yield. Model B (dashed line) is estimated to reproduce the slope of the forward rate regression with maturity $n = 1$.

Figure 3
Properties of the "Mixed" Two-Factor Model



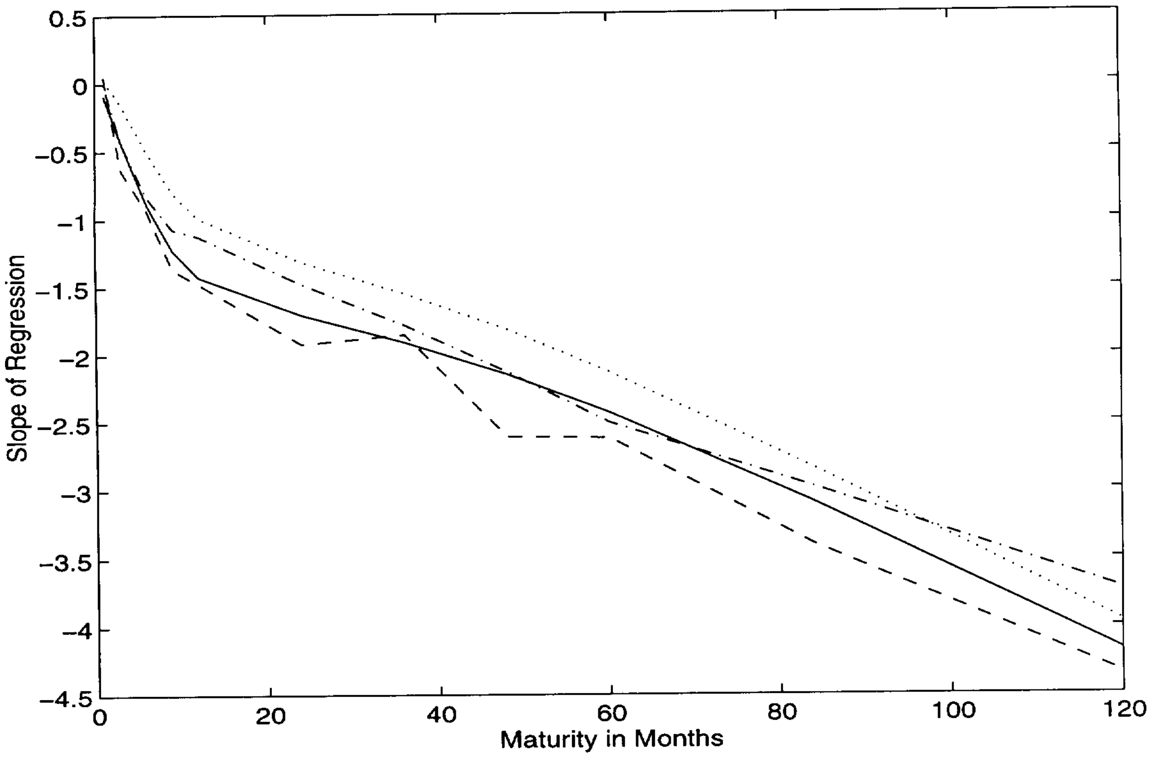
The two panels report forward rate regression slopes and mean yields in the data (the asterisks) and in the two-factor Model H (lines). The solid lines correspond to the point estimates of Model H reported in Table 5. The dashed lines correspond to 5% and 95% quantiles, computed by Monte Carlo as described in the text.

Figure 4
Small Sample Bias in Regression Slopes



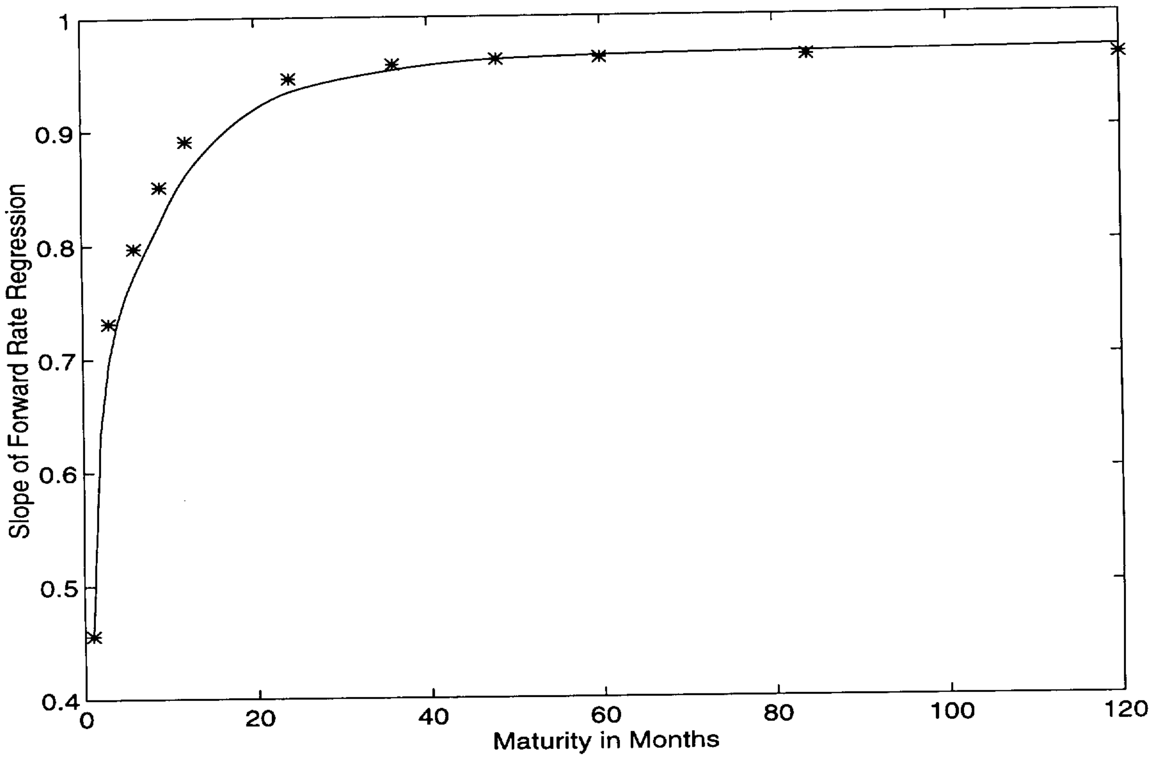
Lines represent small sample bias in slopes c_n of forward rate regressions. The solid line is the estimated bias for the mixed two-factor Model H, computed by Monte Carlo as described in the text. The dashed line is based on equation (9), adapted from Bekaert, Hodrick, and Marshall (1996).

Figure 5
Slopes of Yield Regressions in Four Data Sets



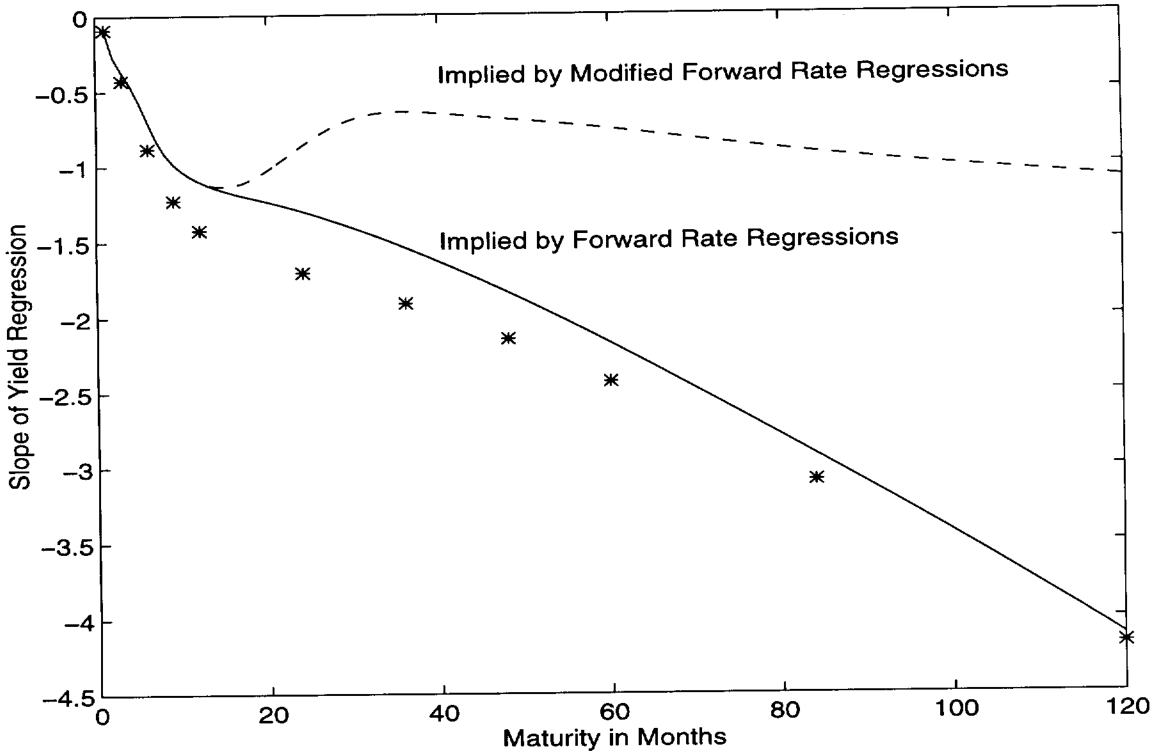
Lines represent the yield regression slopes c_n reported in Table 6 for yields estimated by four different methods. The solid line is based on data estimated by the Smoothed Fama-Bliss method, the dashed line the Unsmoothed Fama-Bliss method, the dash-dotted line the McCulloch Cubic Spline method, and the dotted line the Extended Nelson-Siegel method.

Figure 6
Forward Rate Regressions Implied by Yield Regressions



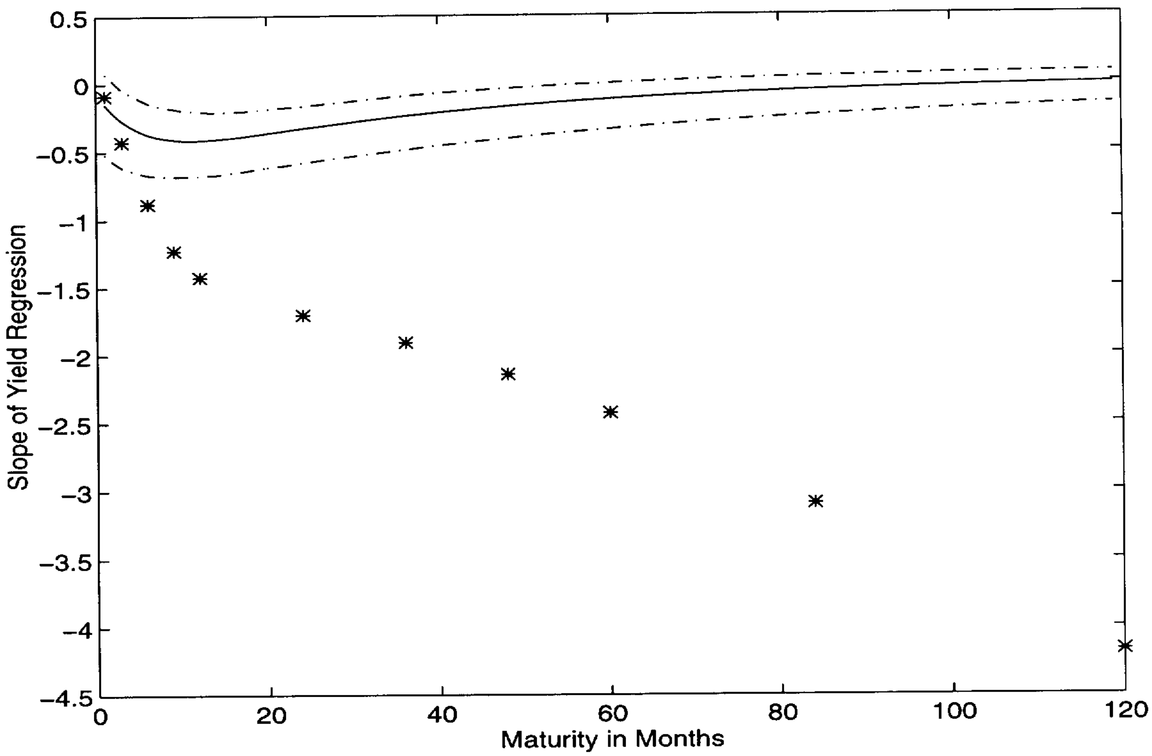
Asterisks represent estimated slopes of forward rate regressions from Table 1 (Smoothed Fama-Bliss data). The line represents slopes implied by a one-factor interpretation of the analogous yield regressions in Table 6.

Figure 7
Yield Regressions Implied by Forward Rate Regressions



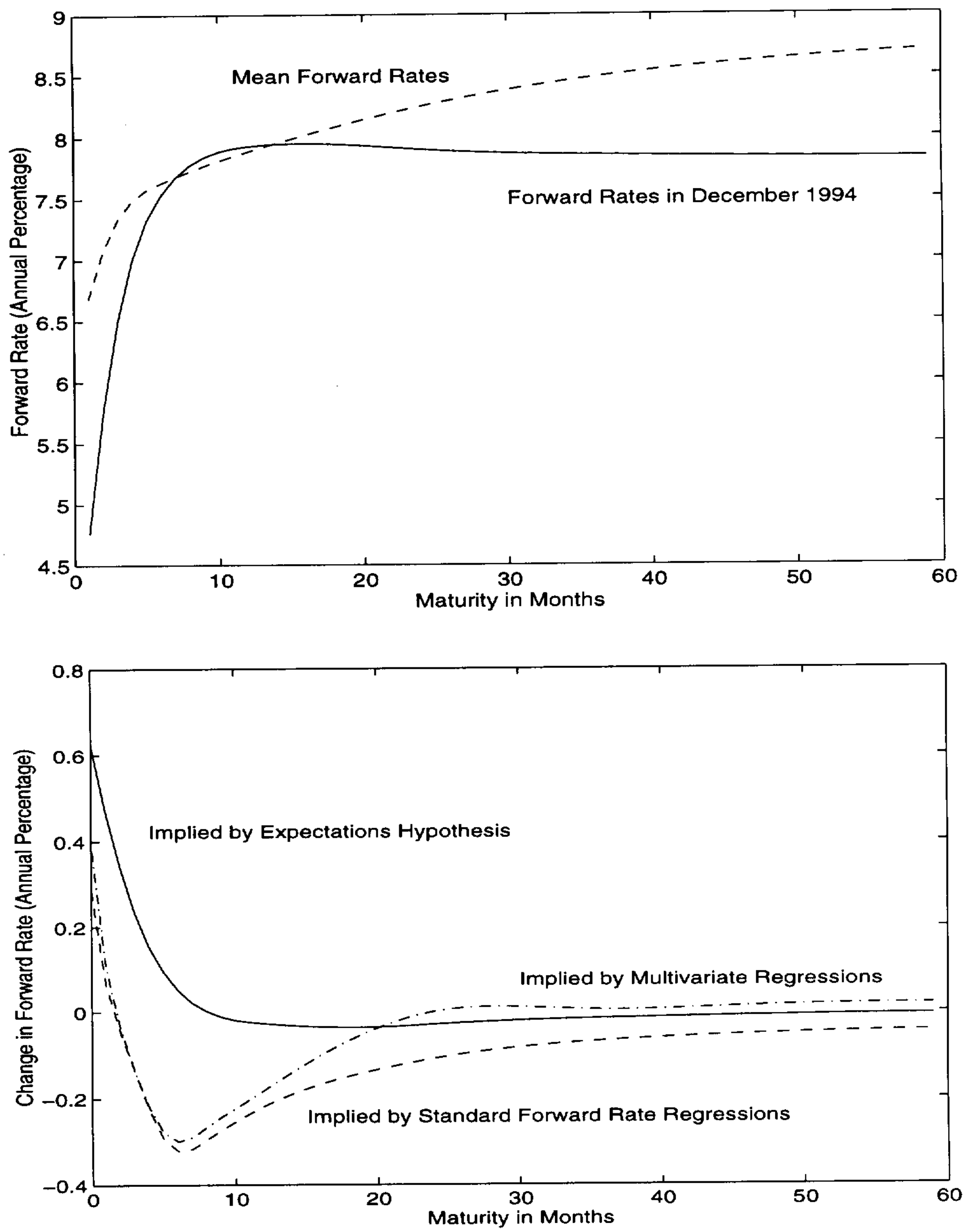
Asterisks represent estimated slopes of yield regressions from Table 6 (Smoothed Fama-Bliss). The solid line represents slopes implied by a one-factor interpretation of the analogous forward rate regressions in Table 1. The dashed line represents implied slopes when forward rate slopes are set equal to one for maturities greater of 24 months or more.

Figure 8
Other Regressions Implied by the "Mixed" Two-Factor Model



Asterisks represent slopes of yield regression slopes from Table 6 (Smoothed Fama-Bliss data). Lines represent regression slopes implied by Model H, a two-factor model with one positive and one negative factor. The dashed lines correspond to 5% and 95% quantiles, computed by Monte Carlo as described in the text. The solid line corresponds to the median.

Figure 9
Forecasts of the Forward Rate Curve



The top panel reports the forward rate curve of December 1994 (solid line) and the mean forward rate curve (dashed line). The bottom panel reports three forecasts of one-month changes in forward rates between December 1994 and January 1995. The forecasts are based, respectively, on the expectations hypothesis (solid line), estimates of equation (7) (dashed line), and estimates of equation (23) (dash-dotted line).