TWO FACTORS ALONG THE YIELD CURVE

by
Frank F. Gong and Eli M. Remolona

Federal Reserve Bank of New York
Research Paper No. 9613

June 1996

Draft, May 1996

Two Factors Along the Yield Curve

FRANK F. GONG and ELI M. REMOLONA[1]

Capital Markets Function
Federal Reserve Bank of New York
New York, NY 10045
Tel: (212)720-6943(Gong)
(212)720-6328(Remolona)
E-mail: frank.gong@frbny.sprint.com
eli.remolona@frbny.sprint.com

Abstract

We estimate two-factor equilibrium models on different parts of the yield curve. In this exploration of the term structure of interest rates, we use two-factor affine yield models as our diagnostic tool. The exercise provides insights on how to reconcile the time-series dynamics of interest rates with the cross-sectional shapes of the term structure and on how movements in the yield curve are related to macroeconomic fundamentals. The evidence favors models in which one factor reverts over time to a time-varying mean. One such model seems adequate to explain three-month to two-year bond yields and another such model to explain two-year to ten-year yields. The models differ because mean reversion is much faster for yields near the short end of the curve than for yields near the long end. Near the short end, the implied factors capture mean reversion in inflation and the Federal Reserve's federal funds target rate. Near the long end, the factors also track the federal funds target but not inflation.

JEL Classification Codes: E43, G12, G13.

Keywords: Term structure, two-factor models, pricing kernel, affine yields, mean reversion, time-varying mean, Kalman filter, economic fundamentals.

[1]We thank David Backus, Ben Bernanke, Arturo Estrella, Richard Cantor, Joe Dziwura, Young Ho Eom, Ethan Harris, Thomas Ho, Tony Rodrigues, and Michael Wickens for helpful discussions. The views expressed in the paper are those of the authors and are not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

Two Factors Along the Yield Curve

FRANK F. GONG and ELI M. REMOLONA

Federal Reserve Bank of New York

EQUILIBRIUM MODELS of the term structure of interest rates face two empirical tasks. The first task is to reconcile the time-series dynamics of interest rates with the cross-sectional shapes of the term structure. The second is to explain movements in the yield curve in terms of observable economic fundamentals. While the literature in economics and finance has made progress on both fronts, thus far advances at one task have hardly been used to further the other task. This paper is an effort to estimate and evaluate models that are consistent with both time-series and cross-section data and to use the estimated models to understand the nature of the fundamentals driving movements in the yield curve.

Theoretical work with equilibrium models, notably by Vasicek (1977) and Cox, Ingersoll, and Ross (1985, hereafter CIR), exploited a no-arbitrage condition to show how the term structure at a moment in time would reflect movements of interest rates over time. In particular, long-maturity yields would depend on market expectations about future interest rates and on risks associated with the volatility of those rates. In the simplest such models, the short-term interest rate is the single factor driving movements in the term structure, and the key parameters are typically the rate by which that short rate returned to a long-run mean and the sensitivity of the rate's volatility to its level.

In reconciling time series data with cross section data on interest rates, the number of factors required for an adequate model is an important issue.

The challenge is to estimate a consistent model with as few factors as possible by relying on discernible regularities in the data. In this regard, one-factor models have not fared well, as Backus and Zin (1994) and Campbell, Lo, and MacKinlay (1994, hereafter CLM) have emphasized.[2] When the parameters are estimated from the time series, the constructed yield curves fail to match the shapes of actual curves, particularly at long maturities.[3] The time-series approach also has difficulty fitting the term structure of volatility, which is critical for pricing fixed-income derivatives. When the parameters are derived from the cross-sectional term structure at a point in time, the estimates suffer from the problem that they inevitably vary over time with shifts in the term structure.[4] Two-factor models hold more promise in fitting the data.[5] Longstaff and Schwartz (1992), for example, propose a model with the short rate and its volatility as the two factors, and the model appears to fit the yield curve up to the five-year maturity.

A separate body of research has demonstrated empirical links between

[2]Campbell, Lo, and Mackinlay state "But in simple term structure models, there also appear to be systematic differences between the parameter values needed to fit cross-section term structure data and the parameter values implied by the time-series behavior of interest rates.

[3]Chan, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1995), Eom (1995), and Stanton (1996) provide alternative estimates of the time-series parameters.

[4]Brown and Dybvig (1986), Ho and Lee (1986), Black, Derman, and Toy (1990), Hull and White (1990), Brown and Schaefer (1991), and Heath, Jarrow, and Morton (1992) show alternative ways of extracting the time-series behavior of rates, particularly volatilities, from the term structure at a moment in time.

[5]The proposed factors differ from model to model: for Brennan and Schwartz (1979) the factors are the short and long rates, for Schaefer and Schwartz (1984) the long rate and the spread between long and short rates, and for Longstaff and Schwartz (1992) the short rate and its volatility. In the present paper, the two-factor models we specify have linear transformations that produce the above models as reduced forms.

the yield curve and observed macroeconomic fundamentals. This research has shown that the yield curve helps predict inflation (Fama 1990, Mishkin 1990, and Engsted 1995), business cycles (Estrella and Hardouvelis 1991 and Estrella and Mishkin 1995), and monetary stance (Rudebusch 1995 and Remolona, Dziwura, and Pedraza 1996). The estimated relationships, however, are based on reduced forms that lack the benefit of structural restrictions imposed by equilibrium models. Some of the studies test the pure expectations hypothesis, which implies a term structure model with constant term premia, but invariably find that the data reject the model.

In this paper, we focus on equilibrium models with only two factors to see how much we can explain with just this number of factors. Instead of trying to fit one model to the whole term structure, however, we find it instructive to try to fit three alternative models, each one to five different parts of the yield curve, using only two maturities at a time. If a two-factor model explained the whole term structure, then we should be able to estimate the model with any two maturities, and the same model should fit different parts of the curve well.[6] The three models we consider include one in which the two factors are additive in the expected log stochastic discount factor as in CLM (1994) and two in which one factor reverts to the other over time but with different shock specifications. We use quarterly data on U.S. zero-coupon yields from 1984 Q1 to 1995 Q1, using five different cross sections of the yield curve, with three-month yields at the short end and ten-year yields at the long end.

We find that the data tend to favor models in which one factor reverts

[6]This exercise is analogous to one conducted by Brown and Weinstein (1983) to test the number of factors in the arbitrage pricing model for the stock market. They divide their sample into several subsets of stocks to see whether they get similar estimates of the risk-free rate and market price of risk.

over time to a factor serving as a time-varying mean. Our results also suggest that explaining the entire term structure would require more than two factors. However, two factors seem adequate to explain wide stretches of the curve. Specifically, one two-factor time-varying-mean model best fits yields of between three-month and two-year maturities, and another such model best fits two-year to ten-year yields. The key difference between the two models is that for yields near the short end of the curve, one factor reverts to the time-varying mean rather quickly, while for yields near the long end, the factor reverts rather slowly.

In our models, the factors are latent variables that serve to forecast stochastic discount factors but which are not directly observable. However, we can estimate the models by Kalman filter and maximum likelihood techniques and then use the estimated models to back out the factors from observed bond yields. To identify the underlying fundamentals, we relate the implied factors to such macroeconomic variables as consumer price inflation, real GDP growth, and the Federal Reserve's federal funds target rate. Using implied factors instead of observed yields allows us to control for time-varying term premia. We find that the factors correlate with fundamentals in ways that are consistent with our two-model division of the term structure. Most strikingly, the factors implied by the shorter-term yields seem to predict future inflation rates and federal funds target rates by capturing mean reverting processes in these fundamentals.

In what follows, we begin by specifying three alternative models of the term structure and deriving the theoretical restrictions necessary for consistency. To keep the restrictions tractable, we take advantage of the recent technology of affine yield models. Then we estimate the relevant model parameters by

using the time series of bond yields of two maturities at a time for different parts of the term structure. The latent-variable structure and well-defined distributional assumptions of the models lead us to employ Kalman filter and maximum likelihood techniques while imposing the restrictions implied by the models. Finally, we back out the implied factors from the estimated models and relate the behavior of these factors to movements in fundamentals.

I. Theory: Affine yield two-factor models

In this paper, we will focus on two-factor equilibrium models. Stambaugh (1988) provides formal support for a two-factor model. He uses a matrix-rank test with Generalized Method of Moments (GMM) applied to monthly data on U.S. Treasury bills of two-month to six-month maturities over the period March 1959 to November 1985 and finds strong evidence against models with a single latent variable. Heston (1992), however, investigates the period from February 1970 to May 1988 and finds little evidence against one-factor models. In our view, the need for at least two factors arises from the apparent failure of one-factor models to reconcile the time series dynamics of interest rates with the cross-sectional shapes of the term structure. As Backus and Zin (1994) have emphasized, the basic difficulty seems to be that the yield curve's steepness at the short end implies rapid mean-reversion by the short rate, while the flatness of the volatility curve requires slow mean reversion.[7] In

[7]Another argument is that one-factor models imply perfect correlation among yields of different maturities, while actual correlations are less than perfect. We do not find this argument to be compelling, because the correlations are after all quite high–in our sample the correlation between the three-month yield and ten-year yield is 0.83.

Appendix C, for example, we try to fit a one-factor model with disappointing results. The estimated persistence of the factor is 0.96 for monthly data. This persistence is not high enough to match long-end volatilities.

To see how far we can push two-factor models, we will specify three alternative models that we believe represent most of the plausible explanations of term-structure movements. If a two-factor model is adequate for the entire term structure, then it should not matter which two maturities we use to estimate the relevant parameters. If the estimates turn out to be different for different parts of the yield curve, then we may need a three-factor model. In this paper, the factors will represent unobservable state variables. Unlike the two factor models in CLM (1994), we allow for correlated factor shocks in two of the three proposed models. Allowing for correlation between the shocks is important because we want to relate them to observable macroeconomic fundamentals. Orthogonalized factors would represent linear combinations of fundamentals.[8] In our first model, the two factors enter additively. In the two other models one factor reverts to the other factor over time.

A. The Pricing Kernel

We rely on a no-arbitrage condition common to intertemporal asset pricing models.[9] In the case of zero-coupon bonds, the price of an n-period bond is

$$P_{nt} = E_t[M_{t+1}P_{n-1,t+1}] \tag{1}$$

where M_{t+1} is the stochastic discount factor. The condition expresses the price of the bond as the expected discounted value of the bond's next-period price.

[8]Such factors would be useful if the purpose is simply the hedging of bond portfolios. Litterman and Scheinkman (1991), for example, propose three such factors.

[9]Singleton (1990) provides a critical survey of these models, particularly their empirical performance.

It rules out arbitrage opportunities by applying the same discount factor to all bonds. We will model P_{nt} by modelling the stochastic process for M_{t+1}, a process called the pricing kernel.[10] Indeed we can solve (1) forward to get $P_{nt} = E_t[M_{t+1}...M_{t+n}]$, which specifies bond prices to be simply functions of the future discount factors. By convention we normalize $P_{0t} = 1$ to ensure the equality of a bond's price at maturity to its par value.

B. Two Factor Affine Yields

We will focus on a class of models known as *affine yield models.* In these models, M_{t+1} is conditionally lognormal, bond prices are jointly lognormal with M_{t+1}, and bond yields are linear in the factors that forecast M_{t+1}. Term structure models that impose the no-arbitrage condition can get very complicated, but affine yield models help maintain tractability. This class of models is fairly general, encompassing those by Vasicek (1977), CIR (1985), Longstaff and Schwartz (1992), CLM (1994), and most equilibrium models of the yield curve. Duffie (1992) provides a detailed discussion of this class of models.

The assumption of joint lognormality allows us to take logs of (1) and write it as

$$p_{nt} = E_t(m_{t+1} + p_{n-1,t+1}) + \frac{1}{2}Var_t(m_{t+1} + p_{n-1,t+1}), \tag{2}$$

where lower-case letters denote logarithms of upper-case letters. Furthermore, if we have two factors, x_{1t}, x_{2t}, that forecast m_{t+1}, an affine yield model can be written as

$$-p_{nt} = A_n + B_{1n}x_{1t} + B_{2n}x_{2t}. \tag{3}$$

[10]The term "pricing kernel" is due to Sargent (1977). In consumption-based equilibrium models, M_{t+1} would represent the marginal rate of substitution between present and next-period consumption (Lucas 1978, for example).

Since the n-period bond yield is $y_{nt} = -p_{nt}/n$, yields will also be linear in the factors. The coeficients A_n, B_{1n}, B_{2n} will depend on the stochastic processes of x_{1t}, x_{2t}, implying restrictions across coefficients for bond prices of different maturities. In practice, specifying A_n, B_{1n}, B_{2n} involves solving (2) based on the stochastic processes of x_{1t}, x_{2t} and verifying that (3) holds.

In the following, we introduce three alternative affine yield two-factor models. The models share a number of characteristics. Factors in each of the models are $AR(1)$ processes. Bond yields then follow an $ARMA(2,1)$ process, because of the yields' affine structure.[11] The conditional variances of bond yields are linear in one of the factors and thus are $ARCH$ processes. Finally, various linear transformations of any of the models will produce the two factor models of Brennan and Schwartz (1979), Schaefer and Schwartz (1992), or Longstaff and Schwartz (1992) as reduced forms, with the short rate and the long rate, the long rate and the spread, or the short rate and its volatility as alternate pairs of reduced-form factors. These reduced forms will not distinguish between the three structural models we specify, but we will show that the data will distinguish sharply between them.

D. Model I: Additive Factors

In the first model, two factors directly affect expectations of the stochastic discount factor for the period right ahead. Specifically, the conditional expectation of the negative of the log stochastic discount factor depends on two factors that enter additively:

$$-m_{t+1} = x_{1t} + x_{2t} + w_{t+1} \tag{4}$$

[11] Engel's (1984) analysis implies that the sum of two $AR(1)$ processes would be $ARMA(2,1)$, unless there are common roots.

where w_{t+1} represents the unexpected change in the log stochastic discount factor and will be related to risk. The shock w_{t+1} has mean zero and a variance that will be specified to depend on the stochastic processes of the two factors x_{1t} and x_{2t}. Each of these factors follows a univariate AR(1) process with heteroscedastic shocks described by a square-root process

$$\begin{aligned} x_{1,t+1} &= (1-\phi_1)\mu + \phi_1 x_{1,t} + x_{2,t}^{0.5} u_{1,t+1} \\ x_{2,t+1} &= (1-\phi_2)\theta + \phi_2 x_{2,t} + x_{2,t}^{0.5} u_{2,t+1} \end{aligned} \tag{5}$$

where $1-\phi_1$ and $1-\phi_2$ are the rates of mean reversion, μ and θ are the long-run means to which the factors revert, and $u_{1,t+1}$ and $u_{2,t+1}$ are shocks with mean zero, volatilities σ_1^2 and σ_2^2 and covariance σ_{12}. It is important to allow correlation between factor shocks, because we intend to relate the factors to fundamentals, which may not be orthogonal.

Moreover, we specify the shock to m_{t+1} to be proportional to the shock to $x_{1,t+1}$, which in turn depends on the level of x_{2t}:

$$w_{t+1} = x_{2,t}^{0.5} \lambda u_{1,t+1} \tag{6}$$

where λ represents the market price of risk. When λ is negative, bond returns are inversely correlated with the stochastic discount factor and risk premia are positive.

The model is similar to CLM (1994) except that we allow the factors to be correlated and both factors to influence the shock to m_{t+1}. A possible interpretation of the factors is that x_{1t} represents the real component of the stochastic discount rate while x_{2t} represents the expected inflation component. The shock to the stochastic discount factor then depends on the level of the inflation component and on the volatility of the real component. This inter-

pretation would be supported by a negative estimate of σ_{12} to correspond to the negative correlation between real interest rates and inflation rates.

The normalization $p_{0t} = 0$ gives us affine-yield coefficients of $A_0 = B_{10} = B_{20} = 0$. We can then derive the one-period yield or short rate as

$$\begin{aligned} y_{1,t} &= -p_{1,t} = -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\ &= x_{1,t} + x_{2,t} - \frac{1}{2}\lambda^2\sigma_1^2 x_{2,t} \end{aligned} \tag{7}$$

which is also linear in the factors, with the coefficients $A_1 = 0$, $B_{11} = 1$, and $B_{21} = 1 - \frac{1}{2}\lambda^2\sigma_1^2$.

We can also verify that the price of an n-period bond is linear in the factors with the coefficients restricted by (see Appendix A)

$$\begin{aligned} A_n &= A_{n-1} + (1-\phi_1)\mu B_{1,n-1} + (1-\phi_2)\theta B_{2,n-1} \\ B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\ B_{2,n} &= 1 + \phi_2 B_{2,n-1} - \frac{1}{2}[(\lambda + B_{1,n-1})^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 \\ &+ 2(\lambda + B_{1,n-1})B_{2,n-1}\sigma_{12}], \end{aligned} \tag{8}$$

The coefficients B_{1n} and B_{2n} are bond price sensitivities to the factors, in much the same way that duration measures bond price sensitivity to the short rate. The coefficient A_n represents the pull of the factors to their means μ and θ.

These recursive equations impose cross-sectional restrictions to be satisfied by eight parameters: $\phi_1, \phi_2, \mu, \theta, \sigma_1, \sigma_2, \sigma_{12}$, and λ. If the two-factor model adequately describes the true pricing kernel process, we should be able to estimate all the parameters using the time series of bond yields of any two maturities.

A model is useful in the pricing of fixed-income options because it can provide a consistent volatility term structure. In the case of our model, the conditional variance of the n-period yield is given by

$$Var_t(y_{n,t+1}) = \frac{1}{n^2}(B_{1n}^2\sigma_1^2 + B_{2n}^2\sigma_2^2 + 2B_{1n}B_{2n}\sigma_{12})x_{2t} \tag{9}$$

We would have a downward-sloping volatility curve given that ϕ_1 and ϕ_2 are both less than unity. Mean reversion by the factors serves to dampen yield volatilities as maturity is lengthened.

The model also allows us to measure term premia. We can derive term premia in the form of the expected excess bond return:

$$\begin{aligned} E_t p_{n-1,t+1} - p_{nt} \quad - \quad & y_{1t} = -\lambda(B_{1,n-1}\sigma_1^2 + B_{2,n-1}\sigma_{12})x_{2t} \\ - \quad & \frac{1}{2}(B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + 2B_{1,n-1}B_{2,n-1}\sigma_{12})x_{2t} \end{aligned} \tag{10}$$

The first term represents a risk premium that depends on the covariance between the stochastic discount factor and bond returns, while the second term represents Jensen's inequality arising from the use of logarithms. Positive term premia require that λ be so negative that

$$-\lambda > \frac{1}{2}\frac{(B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + 2B_{1,n-1}B_{2,n-1}\sigma_{12})}{(B_{1,n-1}\sigma_1^2 + B_{2,n-1}\sigma_{12})}$$

Note also that if $\sigma_2 = 0$, we will have homoscedastic shocks, term premia will be constant, and the pure expectations hypothesis will hold.

E. Model II: a time-varying mean model with both factors heteroskedastic

In our first time-varying mean two-factor model, one factor directly affects expectations of the stochastic discount factor for the immediate period, while the second factor affects expectations of the ultimate destination of the discount factor. Specifically, the model specifies the conditional expectation of

the negative of the log stochastic discount factor to depend directly on one factor, but this factor reverts over time to a second factor:

$$\begin{aligned} -m_{t+1} &= x_t + w_{t+1} \\ x_{t+1} &= (1-\phi_1)\mu_t + \phi_1 x_t + \mu_t^{0.5} u_{1,t+1} \\ \mu_{t+1} &= (1-\phi_2)\theta + \phi_2 \mu_t + \mu_t^{0.5} u_{2,t+1}, \\ w_{t+1} &= \lambda \mu_t^{0.5} u_{2,t+1} \end{aligned} \tag{11}$$

where $1-\phi_1$ and $1-\phi_2$ are rates of mean reversion, but we have a single parameter for the mean, θ. The factor shocks $u_{1,t+1}$ and $u_{2,t+1}$ have mean zero, volatilities σ_1^2 and σ_2^2 and covariance σ_{12}. The conditional expectation of the stochastic discount factor depends only on the first factor, while its shock depends on the second factor.

The time-varying-mean model focuses on distinguishing market developments that affect the current discount rate from those that affect future discount rates, although the model also allows expectations about the distant future to affect risk in the present. The model is similar to that of Balduzzi, Das, and Foresi (1994), who specify a model in which the short rate reverts to a central tendency which itself changes stochastically.

Again we guess that we have an affine yield model, so that the log bond price would have the form $-p_{nt} = A_n + B_{1n}x_t + B_{2n}\mu_t$. The normalization $p_{0t} = 0$ satisfies our guess, because it would give us coefficients of $A_0 = B_{10} = B_{20} = 0$. We can further derive the one-period yield or the short rate as

$$\begin{aligned} y_{1t} &= -p_{1t} = A_1 + B_{1,1}x_t + B_{2,1}\mu_t \\ &= x_t - \frac{1}{2}\lambda^2\sigma_2^2\mu_t \end{aligned} \tag{12}$$

which is also linear in the factors, with the coefficients $A_1 = 0$, $B_{1,1} = 1$, and $B_{2,1} = -\lambda^2\sigma_2^2/2$.

As before, it can be shown that the price of a n-period bond is linear in the factors with the coefficients restricted by (see Appendix A)

$$\begin{aligned} A_n &= A_{n-1} + (1-\phi_2)\theta B_{2,n-1} \\ B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\ B_{2,n} &= \phi_2 B_{2,n-1} + (1-\phi_1)B_{1,n-1} - \frac{1}{2}[(\lambda + B_{2,n-1})^2\sigma_2^2 \\ &+ B_{1,n-1}^2\sigma_1^2 + 2(\lambda + B_{2,n-1})B_{1,n-1}\sigma_{12}], \end{aligned} \tag{13}$$

The coefficients B_{1n} and B_{2n} have the same interpretations as in *Model I*; they measure bond price sensitivities to the factors in much the same way that duration measures bond price sensitivity to the short rate. The coefficient A_n represents the pull of the second factor to its mean θ.

The time-varying mean model is more restrictive than the additive model. We have fewer free parameter here and the recursive equations impose cross-sectional restrictions to be satisfied by seven parameters: $\phi_1, \phi_2, \theta, \sigma_1, \sigma_2, \sigma_{12}$, and λ.

The volatility curve in terms of the conditional variance of the n-period yields in this model is given by

$$Var_t(y_{n,t+1}) = \frac{1}{n^2}(B_{1n}^2\sigma_1^2 + B_{2n}^2\sigma_2^2 + 2B_{1n}B_{2n}\sigma_{12})\mu_t, \tag{14}$$

and again, we would have a downward-sloping volatility curve given that ϕ_1 and ϕ_2 are both less than unity. Mean reversion by the factors serves to dampen yield volatilities as maturity is lengthened.

We derive term premia from expected excess bond returns (Appendix B)

$$\begin{aligned} E_t p_{n-1,t+1} - p_{nt} &- y_{1t} = -\lambda(B_{2,n-1}\sigma_2^2 + B_{1,n-1}\sigma_{12})\mu_t \\ &- \frac{1}{2}(B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + 2B_{1,n-1}B_{2,n-1}\sigma_{12})\mu_t \end{aligned} \tag{15}$$

The first term represents a risk premium that depends on the covariance between the pricing kernel and bond returns, while the second term represents Jensen's inequality arising from the use of logarithms.

Note also that if $\sigma_2 = 0$, we will have homoscedastic shocks, term premia will be constant, and the pure expectations hypothesis will hold.

We derive the expected change in the short rate over n periods as follows (Appendix B),

$$E_t(y_{1,t+n}) - y_{1,t} = \alpha_n + \beta_n(\mu_t - x_t) + \gamma_n \mu_t \tag{16}$$

where

$$\begin{aligned}
\alpha_n &= [1 - \phi_1^n - \frac{1-\phi_1}{\phi_1 - \phi_2}(\phi_1^{n+1} - \phi_2^{n+1}) - \frac{1}{2}\lambda^2\sigma_2^2(1-\phi_2^n)]\theta \\
\beta_n &= 1 - \phi_1^n \\
\gamma_n &= \frac{1-\phi_1}{\phi_1 - \phi_2}(\phi_1^{n+1} - \phi_2^{n+1}) + \frac{1}{2}\lambda^2\sigma_2^2(1-\phi_2^n) - (1-\phi_1^n)
\end{aligned}$$

Note that $\alpha_n = -\gamma_n\theta$, so expected changes in the short rates consist of two mean-reversion components: one is the mean reversion of the first factor x_t to its time-varying mean μ_t, the second is the time-varying mean μ_t reverting to its long-run mean θ.

F. Model III. a time-varying mean model with one homoskedastic factor

In *Model II*, the first factor is closely related to the second factor and both are heteroskedastic with the conditional volatility linked to the time-varying mean. Here we specify a slightly different model in which the two factors are not so linked, with one factor having homoskedastic shocks and the other factor having heteroskedastic shocks. This specification allows more flexibility in modeling term premia. We assume the shocks to the factors are

uncorrelated.

$$\begin{aligned} -m_{t+1} &= x_t + w_{t+1} \\ x_{t+1} &= (1-\phi_1)\mu_t + \phi_1 x_t + u_{1,t+1} \\ \mu_{t+1} &= (1-\phi_2)\theta + \phi_2 \mu_t + \mu_t^{0.5} u_{2,t+1}, \\ w_{t+1} &= \lambda \mu_t^{0.5} u_{2,t+1} \end{aligned} \tag{17}$$

As before, $1-\phi_1$ and $1-\phi_2$ are rates of mean reversion, and we have a single parameter for the mean, θ. The factor shocks $u_{1,t+1}$ and $u_{2,t+1}$ have mean zero, volatilities σ_1^2 and σ_2^2 and zero covariance.

The short rate is the same as in *Model II*

$$y_{1t} = x_t - \frac{1}{2}\lambda^2 \sigma_2^2 \mu_t \tag{18}$$

where $A_1 = 0$, $B_{1,1} = 1$, and $B_{2,1} = -\lambda^2\sigma_2^2/2$.

The price of a n-period bond is similarly linear in the two factors with coefficients given by

$$\begin{aligned} A_n &= A_{n-1} + (1-\phi_2)\theta B_{2,n-1} - \frac{1}{2}B_{1,n-1}^2\sigma_1^2 \\ B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\ B_{2,n} &= \phi_2 B_{2,n-1} + (1-\phi_1)B_{1,n-1} - \frac{1}{2}(\lambda + B_{2,n-1})^2\sigma_2^2 \end{aligned} \tag{19}$$

The term $\frac{1}{2}B_{1,n-1}^2\sigma_1^2$ which was in $B_{2,n}$, as specified in (13), is now in A_n, since the first factor is no longer heteroskedastic. The terms containing the covariance σ_{12} disappear.

The conditional variance of the n-period yield is

$$Var_t(y_{n,t+1}) = \frac{1}{n^2}(B_{1n}^2\sigma_1^2 + B_{2n}^2\sigma_2^2\mu_t) \tag{20}$$

which now contains a constant term and a time-varying term.

We derive term premia from expected excess bond returns (Appendix B)

$$\begin{aligned} E_t p_{n-1,t+1} - p_{nt} - y_{1t} &= -\frac{1}{2}B_{1,n-1}^2\sigma_1^2 - \lambda B_{2,n-1}\sigma_2^2\mu_t \\ &- \frac{1}{2}B_{2,n-1}^2\sigma_2^2\mu_t. \end{aligned} \tag{21}$$

The term premia now contain three components: a constant risk premium term, a time-varying risk premium term, and a Jensen's inequality term.

In this model the expected change in the short rate will be the same as in *Model II*, given by (16), since the drift of short rate process depends on the factors in the same way as in *Model II*.

II. Evidence: Estimating the Parameters

In each of the proposed models, we cannot observe the stochastic discount factor and the underlying two factors directly. All we can observe are the zero-coupon rates on the yield curve. These latent-factor affine models can be cast in the *state space* form and be estimated by the maximum likelihood method combined with the Kalman filter. We depart from the standard application of the Kalman filter by imposing restrictions from our equilibrium models.

The application of the Kalman filter in the estimation of term structure models has been utilized by Jegadeesh and Pennacchi (1996), where they estimate a homoskedastic two-factor term structure model using Eurodollar futures. We use the observed zero-coupon yields to estimate our restricted

latent-variable models, in which the underlying factors are specified to be conditionally heteroskedastic. Heteroskedasticity is important in our model since it leads to time-varying risk premia and, empirically it is well known that interest rates are heteroskedastic. The Kalman filter is the optimal estimation procedure because it is designed to exploit conditional information.

We estimate model parameters using the time series of a pair of yields at a time, imposing on the yield dynamics cross section restrictions implied by the model. If a two-factor model is adequate to characterize the movements of the yield curve, then two yields are enough to back out the factors and it should not matter which two yields are used, as long as the theoretical cross-section restrictions are imposed. By using just two maturities at a time, we will tend to get the models to fit either very tightly or not at all. This helps to sharply differentiate each part of the yield curve from other parts. This approach is analogous to that of Brown and Weinstein (1983), who estimate an arbitrage pricing model using different groups of securities. They argue that if the model is adequate, the factor prices and implied risk-free rate should be the same across the groups.

Recent efforts to reconcile the time series with the cross section have tended to rely on the *GMM* approach, including Longstaff and Schwartz (1992), Gibbons and Ramaswamy (1993), and Backus and Zin (1994). The *GMM* approach places cross-section restrictions on only some of the unconditional moments while the Kalman filter makes use of a richer set of conditional moments. Backus and Zin (1994) place cross-section restrictions on yields up to the 10-year maturity, but the restrictions are placed on only the first moments. Gibbons and Ramaswamy (1993) have cross section restrictions on both the first and the second moments, but use only the yields with maturities up to

one year and avoid using long maturities to fit and test the CIR model. The argument is that the volatility at the long end is mainly due to inflation, and they are only interested in investigating restrictions placed on real returns by the CIR model. In one of the most impressive efforts to date, Longstaff and Schwartz (1992) estimate a model with the short rate and its volatility as the two factors. In testing their overidentifying restrictions, however, they fit a reduced form that takes account of only four out of the ten parameters implied by their structural model.

The use of only some unconditional moment restrictions would present no problem if the purpose is to be able to match the average shape of the term structure. However, since our focus here is to back out the unobserved factors that are consistent with the equilibrium model and relate them to observed economic fundamentals, it is then crucial for the backed-out factor to be both dynamically and cross-sectionally consistent. Cross-section restrictions on only unconditional first moments would not be sufficient. In the maximum likelihood framework, we are able to place cross-section restrictions on the entire distribution (in this case, the first and the second moments) of the yields used in the estimation.

In spirit, our work is close to Backus and Zin (1994) in that we use the observed yields to determine the dynamics of the underlying stochastic discount factor. Our work differs from Backus and Zin (1994) in an important respect: they estimate a reduced form in the sense that they study various ARMA processes for the stochastic discounting factor. We estimate a structural model by specifying the underlying factors that drive the movements of the stochastic discounting factor. An $ARMA(2,1)$ yield process is generated by our two $AR(1)$ factors.

A. Data and Summary Statistics

We obtain end-of-quarter zero-coupon yield data from McCulloch and Kwon (1993) for 1984 Q1 to 1990 Q4 and from the Federal Reserve Bank of New York for 1991 Q1 to 1995 Q1.[12] In the case of the Federal Reserve data, each zero curve is generated by fitting a cubic spline to prices and maturities of about 160 outstanding coupon-bearing U.S. Treasury securities. The securities are limited to off-the-run Treasuries to eliminate the most liquid securities and reduce the possible effect of liquidity premia. Fisher, Nychka, and Zervos (1995) explain the procedure in detail. Summary statistics for the yields with maturities of 1, 2, 4, 8, 20, and 40 quarters for the sample period 84:Q1-95:Q3 are reported in Table 1. The average term-structure is upward sloping, with mean yields ranging from 5.94% to 8.43%. The average term structure of volatility is hump-shaped, with the highest volatility at the 1-2 year maturities. Overall, the volatility curve is very flat. It is interesting that the monthly and quarterly yields across the curve are all very persistent, with first-order monthly auto-correlations at 0.96-0.98, and quarterly first-order autocorrelation at 0.87-0.90. We also conducted the Durbin-Watson test and Ljung-Box Q test to the residual of the AR(2) regression, and the test does not reject the hypothesis that there is no higher order autocorrelation beyond the second order.

We fit each of our models to five different cross sections of the yield curve, using two maturities at a time. These maturity pairs are: (1) the three-month and six-month yields, (2) the six-month and one-year yields, (3) the one-year and two-year yields, (4) the two-year and five-year yields, and (5)

[12]We also examined monthly data, but the estimation would not converge in longer maturity yields. When convergence was achieved, the results were similar to quarterly data.

the five-year and ten-year yields. Table 2 shows the correlations between the yields of different maturities. We see that the zero-coupon yields are all highly correlated, especially between neighboring yields. The correlation between three-month and six-month yields is 0.996, while that correlation between five-year and ten-yields is 0.987. Even the correlation between the two extreme maturities, the three-month and the ten-year yields is 0.828. These correlations will tend to make it difficult to fit the models. Indeed, as we shall see, the models will tend to either fit well or not at all. Nonetheless, we believe that we gain important insights by concentrating on specific parts of the yield curve.

B. Kalman Filtering and Maximum Likelihood Estimation

We now show how to fit each of the two-factor models to U.S. zero-coupon rates data, using a pair of yields at a time.

B1. Model I: The Additive Model

We write the model in the linear state-space form, with the measurement equation

$$\begin{bmatrix} y_{n,t} \\ y_{m,t} \end{bmatrix} = \begin{bmatrix} a_{k_1} \\ a_{k_2} \end{bmatrix} + \begin{bmatrix} b_{1,n} & b_{2,n} \\ b_{1,m} & b_{2,m} \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} \tag{22}$$

where $y_{n,t}$ and $y_{m,t}$ are zero-coupon yields at time t with maturities n and m and v_t is a measurement error assumed to be i.i.d. as

$$v_t \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \delta_1^2 & 0 \\ 0 & \delta_2^2 \end{pmatrix}), \tag{23}$$

and $a_k = A_k/k, b_{1,k} = B_{1,k}/k, b_{2,k} = B_{2,k}/k, k = n, m.$

The transition equation is

$$\begin{bmatrix} x_{1,t+1} \\ x_{2,t+1} \end{bmatrix} = \begin{bmatrix} (1-\phi_1)\mu \\ (1-\phi_2)\theta \end{bmatrix} + \begin{bmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + x_{2,t}^{0.5} \begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \end{bmatrix} \quad (24)$$

with shocks to the state variable X_{t+1} distributed as

$$\begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \end{bmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}) \quad (25)$$

In standard linear state-space models, no restrictions link the measurement equation and the transition equation. This time, however, the measurement equation comes from the transition equation and the no-arbitrage conditions, and the restrictions are given by equation (8).

B2. Models II and III: The Time-Varying Mean Models

For the time-varying mean models, we have the measurement equation

$$\begin{bmatrix} y_{n,t} \\ y_{m,t} \end{bmatrix} = \begin{bmatrix} a_{k_1} \\ a_{k_2} \end{bmatrix} + \begin{bmatrix} b_{1,n} & b_{2,n} \\ b_{1,m} & b_{2,m} \end{bmatrix} \begin{bmatrix} x_t \\ \mu_t \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} \quad (26)$$

where v_t is the measurement error assumed to be i.i.d. as

$$v_t \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \delta_1^2 & 0 \\ 0 & \delta_2^2 \end{pmatrix}), \quad (27)$$

and $a_k = A_k/k, b_{1,k} = B_{1,k}/k, b_{2,k} = B_{2,k}/k, k = n, m.$

The difference between *Model II* and *Model III* is in the transition equation and the cross-section restrictions. In the case of *Model II*, the transition equation is

$$\begin{bmatrix} x_{t+1} \\ \mu_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \\ (1-\phi_2)\theta \end{bmatrix} + \begin{bmatrix} \phi_1 & 1-\phi_1 \\ 0 & \phi_2 \end{bmatrix} \begin{bmatrix} x_t \\ \mu_t \end{bmatrix} + \mu_t^{0.5} \begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \end{bmatrix} \quad (28)$$

with shocks to the state variable X_{t+1} distributed as

$$\begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \end{bmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}). \tag{29}$$

The restrictions are given by (13).

In the case of *Model III*, the transition equation is

$$\begin{bmatrix} x_{t+1} \\ \mu_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \\ (1-\phi_2)\theta \end{bmatrix} + \begin{bmatrix} \phi_1 & 1-\phi_1 \\ 0 & \phi_2 \end{bmatrix} \begin{bmatrix} x_t \\ \mu_t \end{bmatrix} + \begin{bmatrix} u_{1,t+1} \\ \mu_t^{0.5} u_{2,t+1} \end{bmatrix} \tag{30}$$

with shocks to the state variable X_{t+1} distributed as

$$\begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \end{bmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}). \tag{31}$$

The restrictions are given by (19).

B3. Filtering and Estimation

After putting the restrictions into the measurement equations, the preceding models can be estimated by maximum likelihood using the Kalman filter. The althorithm is discussed in Appendix D. For more detailed discussions of the Kalman filtering procedure, see, for example, Hamilton (1994).

C. Estimation Efforts

We attempted to fit each of the three models to five different sections of the yield curve. Table 3 shows the cases where we failed to get convergence and reports the mean log likelihoods for the cases in which convergence was achieved. Surprisingly *Model I* – which is the additive model and the seemingly least restrictive one – would not fit the quarterly data. Convergence is not

achieved even after 2,000 iterations. *Model II* fit the shorter-term yields well, with mean log-likelihood statistics ranging from 0.52, 0.33 and 0.22. However, *Model II* failed to fit the longer maturities. We could not get the program to converge in 2,000 iterations. However, one version of the time-varying mean model, *Model III*, in which the correlation between the shocks to two factors is restricted to zero, did fit the data from the short-end to the long-end. Near the short-end, however, the goodness of fit measured by the mean log-likelihood function shows *Model III* to fit significantly poorer than *Model II* in this part of the curve. All the estimated standard deviations of the measurement error are estimated to be zero except in two occasions. These standard deviations indicate that when a model fits, it tends to fit extremely well.

The data apparently favor two model specifications, one for yields near the short end – from three-month to two-year maturities – and one for yields near the long-end – from two-year to ten-year maturities. Both models specify the second factor to be a time-varying mean to which the first factor reverts. The two models differ, however, in the specification of the risk premium and in the correlation between the shocks to the two factors. In the model that fits shorter-term yields best, both factors have heteroscedastic shocks that are related to the square-root of the second factor. The shocks between factors are correlated. In the model that fits longer-term yields best, one factor has homoscedastic shocks and the shocks between factors are uncorrelated. This induces term premia with constant as well as time-varying components..

Our efforts to fit the models to monthly data were not nearly as successful, especially for the longer maturities. The reason, as we shall see, is that for longer-term yields the first factor follows a near-unit root process. This rather high degree of persistence is apparently is hard to estimate at the monthly

frequency.

III. Parameter Estimates and Interpretation

Fitting two-factor models to different parts of the yield curve sheds light on the issue of the number of factors required for an adequate term structure model. It is significant that a single two-factor model does not fit the entire yield curve well, because this suggests a minimum requirement of three factors. However, it is also important to know whether two separate two-factor models accomplish the task, because this narrows down the requirements, giving us greater hope that we may someday fit to the whole curve a single appropriately specified model with no more than three or four factors.[13]

A. Parameter Estimates

The pattern of parameter estimates suggests that one model would explain the three different sections of the yield curve near the short end and another model would explain the two sections near the long end. We report in Table 4 the estimated parameters for the two time-varying mean models that fit each section of the yield curve best. The parameter estimates across the yield curve from three-month yields to two-year yields are remarkably similar. This result is remarkable because it is based on estimates that rely on two neighboring maturities at a time, which would tend to produce very tight but very different estimates. The similarity between parameter estimates across the curve from the two-year to ten-year maturities is also impressive.

[13]Our own efforts to fit such a model have so far been unsuccessful. We consider all our efforts to be an exercise in data compression in the sense of Sims (1996).

Recall that $1-\phi_1$ and $1-\phi_2$ are the rates of mean-reversion of the first and the second factor respectively. We see that the time-varying mean implied at different part of the yield curve is always relatively persistent. For the short yields, however, the first factor reverts to the time-varying mean rather quickly, with the persistence parameter ranging from 0.50 to 0.80.[14] Note that the yields themselves are much more persistent than the first factor. By specifying a time-varying mean process, the factor that tracks the short rate closely turns out to actually revert to the varying mean quickly. For the longer yields, the first factor is extremely persistent. This makes perfect sense, because the factor should represent short rate movements that are so long lasting that they are reflected in long-term yields.

Note that correlations between the shocks of the two factors are significantly positive in the 3-month to two-year range. From Table 4, we see that the model with correlated factors has a better fit in the 3-month to 2-year part of the yield curve. This highlights the importance of allowing correlated factors near the short end. However, near the long-end, only *Model III* fits the data , and the model succeeds by assuming the correlation between the shocks to be zero.

As expected, the price of risk λ is estimated to be negative. At the short end, it ranges from -1.42 to -2.78. At the long-end, it is -1.04 for 2-5 year range and -5.84 for 5-10 year range. Since the shock to the discount factor is attached to the shock of the second factor, a measure of the size of the risk is $\lambda\sigma_2$. We see that at the two long-end intervals, the numbers are of the same order.

[14]One may compare this with the monthly persistence parameter of 0.35 for the short rate as reported by Gibbons and Ramaswamy (1993).

The estimates show that a single two-factor model does not explain the entire yield curve. The estimates also suggest that we can divide the yield curve into two parts, so that one two-factor model fits one part and another two-factor model fits the other part.

B. Adequacy of a Two-Model Explanation

To assess the adequacy of a two-model explanation of the term structure, the similarity of individual parameter estimates is less important than the combined effects of a given set of estimates, particularly because the parameters are related in a highly nonlinear way. We look at these combined effects by heuristically comparing the yield curves and volatility curves implied by the parameters and by analyzing the comovements of the factors derived from the estimated models.

The yield curves implied by the different estimates appear to sort themselves in a way that corresponds to the division of the yield curve by the two successful model specifications. Figure 1 plots the actual average yield curve for the sample period and the yield curves derived from the unconditional means of yields in the estimated models. Not surprisingly, the estimated curves fit best the parts of the actual curve where they were estimated. Models with strong mean reversion tend to fit the short end, while models with weak mean reversion tend to fit the long end. None of the estimated curves fits well at both the short and long ends, indicating the difficulty of explaining the whole curve with a single two-factor model.[15] At the same time, two models – one for yields near the short end and one for yields near the long end – might

[15]Litterman and Scheinkman (1991) extract three orthogonal latent factors but do not seem to impose the no-arbitrage conditions in their extraction. Their factors seem to represent the level, slope and curvature of the yield curve.

suffice.

The implied volatility curves lead to similar conclusions. Figure 2 plots the actual average volatility curve and the estimated curves based on unconditional yield volatilities. As before, the estimated curves tend to fit most closely the parts of the curve where they were estimated. It is remarkable that the three curves estimated from the shorter-term yields bunch together and the two curves from the longer-term yields bunch together, suggesting a two-model explanation of the volatility curve.

The pattern of correlations among the factors backed out from the different model estimates also suggests a two-model explanation of the term structure. Table 5 reports the correlations among these factors. In Panel (a), the correlations among the various estimates of the first factor – which represents the market's expectation of the next-quarter's stochastic discount factor – are uniformly strong, thus not helping to distinguish between models. In Panel (b), however, the correlations among the various estimates of the time-varying-mean factor are much stronger when the factors belong to a particular group than when the factors belong to different groups. The correlations between the implied time-varying means drawn from three-month to two-year yields range from 0.93 to 0.99, while the correlation between the two estimates drawn from the longer-term yields is 0.87. Outside these groups the correlations are considerably weaker, ranging from 0.14 to 0.65.

IV. Relationship to Fundamentals

A. Fundamentals and Hypotheses

To understand the nature of the economic fundamentals that drive movements in the term structure, we now relate the factors implied by our models to consumer price inflation, real GDP growth, and the Federal Reserve's federal funds target rate. A large literature has demonstrated the yield curve's power to predict such macroeconomic variables. Fama (1990), Mishkin (1990), and Engsted (1995) show that spreads between long and short rates contain information about future inflation rates. Similarly, Estrella and Hardouvelis (1991) and Estrella and Mishkin (1995) show that yield spreads can predict real activity growth and recessions. Rudebusch (1995) suggests that the ability of yield spreads to predict short rates at near horizons is due to the predictability of the Federal Reserve's actions to raise or lower its federal funds target rate. Remolona, Dziwura, and Pedraza (1996) show that forward rates reflect the anticipation of such monetary stance. In this paper, we control for time-varying term premia by using implied factors instead of observed bond yields to predict fundamentals.[16]

An important point of our analysis is to distinguish between two competing hypotheses about the yield curve's predictive power. Yields may predict fundamentals because market participants have horizon-specific information about the future. In this case, yields with maturities that match the forecasting horizon would perform better than other yields. Alternatively, yields may predict fundamentals because the factors reflect regularities in the time-series

[16]The federal funds rate is an overnight rate. Bernanke and Blinder (1992) consider the rate to be a good measure of monetary stance and a good prediction of real activity. In this paper, we use the Federal Reserve's target rate instead of actual market rate.

dynamics of fundamentals. Backus and Zin (1993), for example, suggest that interest rates follow a long-memory process to reflect such a process in inflation rates. In this case, factors derived from yields that match the forecasting horizon would have little advantage over other factors. Moreover, if mean reversion is an important feature of a fundamental's dynamics, the factors' predictive power would tend to improve with longer forecasting horizons.[17]

B. Contemporaneous correlations

Before turning to the predictive power of the derived factors, we examine contemporaneous correlations between the factors and observed macroeconomic fundamentals. Table 6 report the quarterly correlations between the factors and core CPI inflation rates, real GDP growth rates, and the end-of-quarter federal funds target rate. By themselves, the correlations do not tell us what fundamentals the factors represent. The correlations are useful because they suggest which macroeconomic variables provide current information that influences factor movements. A factor representing expected inflation, for example, may actually be correlated with real GDP growth, which may be perceived as a leading indicator of inflation. Such informational effects may serve to differentiate among factors estimated from different parts of the yield curve.

The correlations involving the first factor, which represents the market's expectation of the next-quarter's stochastic discount factor, suggest similar informational effects of macroeconomic variables across estimates derived from different parts of the yield curve. The strongest correlations reported in Table 8 are those between the factor and the federal funds target, and these correlations

[17]We owe to Fama and Bliss (1987) the idea that mean reversion would make longer horizons more predictable.

are uniformly strong whether the factor is derived from yields near the short end or from yields near the long end. The factor is less strongly correlated with core CPI inflation but the correlations are nonetheless consistently significant across estimates from different parts of the yield curve. None of these implied factors shows significant correlation with real GDP growth.

The correlations involving the other factor, which represents the market's expectation of the stochastic discount factor's future destination, point to a notable difference in the way current macroeconomic variables influence different parts of the yield curve. In Table 6*b*, when this time-varying-mean factor is derived from shorter-term yields, the factor is strongly correlated with the federal funds target rate, less strongly correlated with inflation, and uncorrelated with real GDP growth. In contrast, when the factor is derived from longer-term yields, the factor is weakly correlated with real GDP growth but uncorrelated with either the federal funds target or inflation. This pattern of correlations is consistent with the idea that movements in the term structure can be explained by two models, one for shorter-term yields and one for longer-term yields, with the two-year yield as the rough dividing line.

C. Predictive Power of Mean Reversion

To measure the power of the implied factors to predict fundamentals, we rely on the equation we derived for the conditional expectation of the change in the short rate over n periods. Recall from (26) that the equation is $E_t y_{1,t+n} - y_{1t} = \alpha_n + \beta_n(\mu_t - x_t) + \gamma_n \mu_t$. The term involving $\mu_t - x_t$ represents the effect of the reversion of the first factor to the time-varying mean, while the term involving μ_t represents the effect of the reversion of the time-varying mean to its own long-run mean. As shown in (26), the coefficients α_n, β_n, and γ_n are determined by the parameters $\phi_1, \phi_2, \theta, \sigma_2$ and λ. We can tell that the equation

excludes term premia by noting that λ enters only in squared form as a result of Jensen's inequality. We analyze regressions in which the change in the short rate is replaced by the change in the macroeconomic fundamental and in which $\mu_t - x_t$ and μ_t are the explanatory variables. We would not expect the regression coefficients to precisely match the theoretical coefficients α_n, β_n, and γ_n because of the replacement of the left-hand-side variable. Nonetheless, if the factors reflect mean reversion in macroeconomic fundamentals, the regressions would allow us to predict the fundamentals without the confounding effects of time-varying term premia.

The regression results provide striking evidence that the implied factors from shorter-term yields capture a mean reversion process in core CPI inflation. Table *7a* reports the results of regressions of changes in the inflation rate over varying time horizons on $\mu_t - x_t$ and μ_t as variously estimated from different parts of the yield curve. The R-squared statistics show that the factors drawn from the shorter-term yields have significant predictive power with regard to changes in inflation rates over eight to 12 quarters, while the factors drawn from longer-term yields have little predictive power over these horizons. There appears to be no advantage in the predictive power of factors drawn from yields with maturities matching the forecasting horizon. The factors drawn from one-year and two-year yields, for example, fail to outperform the factors drawn from three-month and six-month yields or from six-month and one-year yields at predicting the change in the inflation rate over eight quarters. Hence, the predictive power of yields does not arise from horizon-specific information about the future. Moreover, the predictive power of the factors drawn from yields with up to two-years in maturity improves as the forecasting horizon is lengthened up to at least 10 quarters. These factors seem to derive their predictive power by reflecting a process in which the inflation rate reverts to

a future mean rate over 10 to 12 quarters.

The regression results show some predictive power with regard to real GDP growth but apparently not because of mean reversion. Table 7*b* reports these results. The factors drawn from two-year and five-year yields show significant predictive power for real activity over six and eight quarters. The predictive power, however, does not improve with the forecasting horizon.

The predictive power with regard to the federal funds target is impressive for factors derived from any part of the yield curve and for almost any forecasting horizon. The pattern suggests that mean reversion may also be an important feature of this fundamental. Unlike the case of inflation, however, the pattern is not entirely consistent with our two-model explanation. Table 7*c* reports R-squared statistics that improve with the time horizon up to eight quarters for the factors derived from the first two pairs of shorter-term yields. However, the predictive performance continues to rise beyond this horizon for the factors derived from one-year and two-year yields as it does for the factors derived from two-year and five-year yields. The pattern might be consistent with three models instead of two, with a model for the medium-term range of one-year to five-year yields as well as for the short end and long end.

D. Interpretation

The implied factors tell an interesting story about these fundamentals. The factors depict a market that behaves largely to anticipate the Federal Reserve's actions as reflected in the federal funds target rate. In this market, the federal funds target rate is the primary determinant of short rates. At the same time, market participants view the Federal Reserve as having a near-term target associated with inflation and a long-term target associated with

unidentified fundamentals about which real activity growth seems to provide information. When the Fed's actions are related to its near-term target, the market believes the actions are likely to be reversed in about two years, and the effects on interest rates are not reflected in yields on bonds with more than two years in maturity. When the actions are related to the long-term target, the market believes the actions are likely to be sustained, and the effects are so long-lasting that they affect yields on bonds with up to ten years in maturity.

V. Conclusion

This paper is largely an effort to explore the term structure of interest rates using two-factor affine yield models as our diagnostic tool. We try to fit not one model but three, and we try to fit each model not once but five times by dividing the yield curve into five separate cross sections. The idea is that if a two-factor model explained the whole term structure, then the same model should fit different parts of the curve well. In estimating the models, we use only two yield maturities at a time, the minimum number of maturities needed, to make it hard to get the same model to fit different parts of the yield curve.

The results of the exercise suggest that a single two-factor model would have difficulty explaining the whole term structure. Remarkably, however, the results also suggest that we can divide the yield curve into two parts, so that one two-factor model fits the shorter yield maturities and another two-factor model the longer maturities, with the two-year yield as a rough dividing line. For either part of the curve, the model that fits best is a time-varying mean specification where one factor reverts to the other over time. The key difference between the two models is that one factor reverts to the time-varying mean rather quickly in the model for shorter-term yields while the factor reverts rather slowly in the model for longer-term yields. The two models also differ in the behavior of term premia and the correlation between shocks of the two factors.

As specified, the models represent the pricing kernels used to price bonds of different maturities. That we seem to get two pricing kernels does not mean the presence of arbitrage opportunities, say between short-maturity and

long-maturity bonds. Rather the results suggest that the true pricing kernel is probably more complicated than the ones we specified, and it probably is driven by at least three factors. Nonetheless if the purpose is to price a bond relative to other bonds, a two-factor pricing kernel may be adequate if the bonds are limited either to three-month to two-year maturities or to two-year to ten-year maturities.

While the factors are specified to be unobserved latent variables, we are able to extract them from the estimated models and relate them to observed macroeconomic fundamentals. In general, the factors seem to relate to fundamentals in ways that are consistent with our two-pricing-kernel explanation of the term structure. Most strikingly, we find that the factors implied by the shorter term yields successfully predict future inflation rates and future Federal Reserve targets for the federal funds rate, and the factors succeed largely by capturing mean reverting processes in these fundamentals. The factors implied by the longer term yields also seem to reflect a mean reverting process in the federal funds target but not in inflation or in real activity growth.

Appendix A

A1. Model I: Recursive Restrictions

We start with the general pricing equation:

$$p_{nt} = E_t(m_{t+1} + p_{n-1,t+1}) + \frac{1}{2}Var_t(m_{t+1} + p_{n-1,t+1})$$

The short rate is derived by setting $p_{0,t} = 0$:

$$\begin{aligned} y_{1t} = -p_{1t} &= -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\ &= x_{1,t} + x_{2,t} - \frac{1}{2}\lambda^2\sigma_1^2 x_{2,t}, \end{aligned}$$

showing the short rate to be linear in the factors.

Now we guess that the price of an n-period bond is affine:

$$-p_{n,t} = A_n + B_{1,n}x_{1,t} + B_{2,n}x_{2,t}.$$

We verify that there exist A_n, $B_{1,n}$, and $B_{2,n}$ that satisfy the general pricing equation:

$$\begin{aligned} E_t(m_{t+1} + p_{n-1,t+1}) &= -A_{n-1} - (1-\phi_1)\mu B_{1,n-1} - (1-\phi_2)\theta B_{2,n-1} \\ &- (1 + \phi_1 B_{1,n-1})x_{1,t} - (1 + \phi_2 B_{2,n-1})x_{2,t} \\ Var_t(m_{t+1} + p_{n-1,t+1}) &= [(\lambda + B_{1,n-1})^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 \\ &+ 2(\lambda + B_{1,n-1})B_{2,n-1}\sigma_{12}]x_{2t} \end{aligned}$$

Thus

$$\begin{aligned} -p_{n,t} &= A_n + B_{1,n}x_{1,t} + B_{2,n}x_{2,t} \\ &= A_{n-1} + (1-\phi_1)\mu B_{1,n-1} + (1-\phi_2)\theta B_{2,n-1} \end{aligned}$$

$$
\begin{aligned}
&+ \quad (1+\phi_1 B_{1,n-1})x_{1,t} \\
&+ \quad \{1+\phi_2 B_{2,n-1} - \frac{1}{2}[(\lambda + B_{1,n-1})^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + 2(\lambda + B_{1,n-1})B_{2,n-1}\sigma_{12}]\}x_{2t}
\end{aligned}
$$

Matching coefficient on both sides, we have

$$
\begin{aligned}
A_n &= A_{n-1} + (1-\phi_1)\mu B_{1,n-1} + (1-\phi_2)\theta B_{2,n-1} \\
B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\
B_{2,n} &= 1 + \phi_2 B_{2,n-1} - \frac{1}{2}[(\lambda + B_{1,n-1})^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 \\
&+ 2(\lambda + B_{1,n-1})B_{2,n-1}\sigma_{12}].
\end{aligned}
$$

A2. The Time-Varying Mean Model II: Recursive Restrictions

The short rate is now

$$
\begin{aligned}
r_t = -p_{1t} &= -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\
&= x_t - \frac{1}{2}\lambda^2\sigma_2^2\mu_t,
\end{aligned}
$$

and it is linear in the factors as before.

If the price of a n-period bond is affine, we have:

$$
\begin{aligned}
E_t(m_{t+1} + p_{n-1,t+1}) &= -A_{n-1} - (1-\phi_2)\theta B_{2,n-1} \\
&- (1+\phi_1 B_{1,n-1})x_t \\
&- [(1-\phi_1)B_{1,n-1} + \phi_2 B_{2,n-1}]\mu_t \\
Var_t(m_{t+1} + p_{n-1,t+1}) &= [(\lambda + B_{2,n-1})^2\sigma_2^2 + B_{1,n-1}^2\sigma_1^2 \\
&+ 2(\lambda + B_{2,n-1})B_{1,n-1}\sigma_{12}]\mu_t
\end{aligned}
$$

where

$$
-p_{n,t} = A_n + B_{1,n}x_t + B_{2,n}\mu_t
$$

$$
\begin{aligned}
&= A_{n-1} + (1-\phi_2)\theta B_{2,n-1} \\
&+ (1+\phi_1 B_{1,n-1})x_t \\
&+ \{(1-\phi_1)B_{1,n-1} + \phi_2 B_{2,n-1} \\
&- \frac{1}{2}[B_{1,n-1}^2\sigma_1^2 + (\lambda + B_{2,n-1})^2\sigma_2^2 + 2(\lambda + B_{2,n-1})B_{1,n-1}\sigma_{12}]\}\mu_t
\end{aligned}
$$

Matching coefficient on both sides, we have

$$
\begin{aligned}
A_n &= A_{n-1} + (1-\phi_2)\theta B_{2,n-1} \\
B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\
B_{2,n} &= \phi_2 B_{2,n-1} + (1-\phi_1)B_{1,n-1} - \frac{1}{2}[(\lambda + B_{2,n-1})^2\sigma_2^2 \\
&+ B_{1,n-1}^2\sigma_1^2 + 2(\lambda + B_{2,n-1})B_{1,n-1}\sigma_{12}].
\end{aligned}
$$

Appendix B

B1. Model II: The Term Premia

Term premia can be derived from the expected excess bond return over the short rate:

$$\begin{aligned}
E_t p_{n-1,t+1} - p_{n,t} - y_{1t} &= -A_{n-1} - B_{1,n-1}E_t x_{t+1} - B_{2,n-1}E_t\mu_{t+1} \\
&+ A_n + B_{1,n}x_t + B_{2,n}\mu_t - x_t + \frac{1}{2}\lambda^2\sigma_2^2\mu_t \\
&= (A_n - A_{n-1}) + (B_{1,n} - 1 - \phi_1 B_{1,n-1})x_t \\
&+ (B_{2n} + \frac{1}{2}\lambda^2\sigma_2^2 - \phi_2 B_{2,n-1} - (1-\phi_1)B_{1,n-1})\mu_t \\
&= -\lambda(B_{2,n-1}\sigma_2^2 + B_{1,n-1}\sigma_{12})\mu_t \\
&- \frac{1}{2}(B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + 2B_{1,n-1}B_{2,n-1}\sigma_{12})\mu_t
\end{aligned}$$

B2. Model II: Expected Change in the Short Rate

The conditional expectation of the short rate n periods in the future is

$$\begin{aligned}
E_t y_{1,t+n} &= E_t x_{t+n} - \frac{1}{2}\lambda^2\sigma_2^2 E_t\mu_{t+n} \\
E_t\begin{bmatrix} x_{t+1}-\theta \\ \mu_{t+1}-\theta \end{bmatrix} &= \begin{bmatrix} \phi_1 & 1-\phi_1 \\ 0 & \phi_2 \end{bmatrix}\begin{bmatrix} x_t-\theta \\ \mu_t-\theta \end{bmatrix} \\
E_t\begin{bmatrix} x_{t+n}-\theta \\ \mu_{t+n}-\theta \end{bmatrix} &= \begin{bmatrix} \phi_1 & 1-\phi_1 \\ 0 & \phi_2 \end{bmatrix}^n\begin{bmatrix} x_t-\theta \\ \mu_t-\theta \end{bmatrix} \\
&= \begin{bmatrix} \phi_1^n & \frac{1-\phi_1}{\phi_1-\phi_2}(\phi_1^{n+1} - \phi_2^{n+1}) \\ 0 & \phi_2^n \end{bmatrix}\begin{bmatrix} x_t-\theta \\ \mu_t-\theta \end{bmatrix}
\end{aligned}$$

We can then write

$$E_t y_{1,t+n} - y_{1,t} = \alpha_n + \beta_n(\mu_t - x_t) + \gamma_n\mu_t$$

where

$$\begin{aligned}
\alpha_n &= [1-\phi_1^n - \frac{1-\phi_1}{\phi_1-\phi_2}(\phi_1^{n+1}-\phi_2^{n+1}) - \frac{1}{2}\lambda^2\sigma_2^2(1-\phi_2^n)]\theta \\
\beta_n &= 1-\phi_1^n \\
\gamma_n &= \frac{1-\phi_1}{\phi_1-\phi_2}(\phi_1^{n+1}-\phi_2^{n+1}) + \frac{1}{2}\lambda^2\sigma_2^2(1-\phi_2^n) - (1-\phi_1^n)
\end{aligned}$$

B2. Model III: The Term Premia

$$\begin{aligned}
E_t p_{n-1,t+1} - p_{n,t} - y_{1t} &= -A_{n-1} - B_{1,n-1}E_t x_{t+1} - B_{2,n-1}E_t\mu_{t+1} \\
&+ A_n + B_{1,n}x_t + B_{2,n}\mu_t - x_t + \frac{1}{2}\lambda^2\sigma_2^2\mu_t \\
&= (A_n - A_{n-1}) + (B_{1,n} - 1 - \phi_1 B_{1,n-1})x_t \\
&+ (B_{2n} + \frac{1}{2}\lambda^2\sigma_2^2 - \phi_2 B_{2,n-1} - (1-\phi_1)B_{1,n-1})\mu_t \\
&= -\lambda B_{2,n-1}\sigma_2^2\mu_t - \frac{1}{2}B_{2,n-1}^2\sigma_2^2\mu_t \\
&- \frac{1}{2}B_{1,n-1}^2\sigma_1^2.
\end{aligned}$$

Appendix C

A Homoskedastic Single-factor Model: An example

To illustrate how the time-series dynamics and cross section restrictions of the yields are related, we use the statistics in Table 1 to calibrate a simple term-structure model: In our two-factor model (4)-(7), if we set $\sigma_2 = 0, \phi_2 = 0$, the model reduces to a homoskedastic single-factor model as in CLM (1994). That is

$$\begin{aligned} -m_{t+1} &= x_t + w_{t+1} \\ x_{t+1} &= (1-\phi)\mu + \phi x_t + u_{t+1} \\ w_{t+1} &= \lambda u_{t+1} \end{aligned}$$

Parameters of the model, ϕ, σ^2, μ, and λ, can be estimated by

$$Corr[y_{1t}, y_{1,t-1}] = \phi \tag{32}$$

$$Var[y_{1t}] = \frac{\sigma^2}{1-\phi^2} \tag{33}$$

$$\begin{aligned} E[r_{n,t+1} - y_{1t}] &= E(p_{n-1,t+1} - p_{n,t} + p_{1,t}) \\ &= E[ny_{n,t} - (n-1)y_{n-1,t+1} - y_{1t}] \\ &= -(\frac{1-\phi^{n-1}}{1-\phi})^2 \frac{\sigma^2}{2} - \lambda\sigma^2 \frac{1-\phi^{n-1}}{1-\phi} \end{aligned} \tag{34}$$

$$E[y_{1t}] = \mu - \lambda^2 \frac{\sigma^2}{2} \tag{35}$$

Thus, the first auto-correlation of short rates (the one-month yields) determines the persistence parameter $\phi = 0.96$; with the volatility of the short rates and (33), we have $\sigma^2 = (1 - 0.96^2) * 1.93^2 = 0.29$.

Note that in this single-factor model, the term structure of volatility depends on parameters ϕ and σ^2 only, since

$$y_{nt} = \frac{1}{n}(A_n + B_n x_t)$$

$$Var(y_{nt}) = \frac{1}{n^2} B_n^2 var(x_t) = \frac{1}{n^2} (\frac{1-\phi^{n-1}}{1-\phi})^2 \frac{\sigma^2}{1-\phi^2}.$$

That is, the dynamics of the short rates (its first-order autocorrelation and volatility) determines the whole term structure of volatility.

The implied and actual volatility term structure (84:1-95:3) are shown in Figure A1. We see that, the model estimated from the short rate dynamics fit poorly the actual term structure of volatility.

To estimate the parameters λ and μ, in addition to the short rates, we have to use the average yield of another maturity. Using the average 10-year and 1-month yields, we find $\lambda = -12.75$ and $\mu = 29.51$; using the average yields of 3-month and 1-month, we find $\lambda = -1.53$ and $\mu = 6.28$. The implied term structure of zero-coupon yields using $\phi = 0.96, \sigma^2 = 0.29, \lambda = -1.53, \mu = 6.28$ is shown in Figure A2. We see that it fits the term structure at the short end (up to 3 months), but has a poor fit for longer maturities.

This example highlights the need for a model with more than one factor to reconcile time-series dynamics with cross-section restrictions of the term structure. Also, cross-section restrictions beyond the first moment are needed to fit the term structure of volatility.

Appendix D: The Kalman Filter Algorithm

For the state-space models in section II, the measurement and transition equations can be written in the following matrix form:

Measurement Equation:

$$y_t = A + BX_t + v_t$$

where $v_t \sim N(0, R)$.

Transition Equation:

$$X_{t+1} = C + FX_t + u_{t+1} \tag{36}$$

where $u_{t+1|t} \sim N(0, Q_t)$.

The Kalman filter algorithm of this state-space model is the following:

1. Initialize the state-vector S_t:

The recursion begins with a guess $S_{1|0}$, usually given by

$$\hat{S}_{1|0} = E(S_1). \tag{37}$$

The associated MSE is

$$\begin{aligned} P_{1|0} &\equiv E[(S_1 - \hat{S}_{1|0})(S_1 - \hat{S}_{1|0})'] \\ &= Var(S_1). \end{aligned}$$

The initial state S_1 is assumed to be $N(\hat{S}_{1|0}, P_{1|0})$.

2. Forecast y_t:

Let I_t denote the information set at time t. Then

$$\begin{aligned} \hat{y}_{t|t-1} &= A + BE[S_t|I_{t-1}] \\ &= A + B\hat{S}_{t|t-1}. \end{aligned} \tag{38}$$

The forcasting MSE is

$$E[(y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1})'] = BP_{t|t-1}B' + R \tag{39}$$

3. Update the inference about S_t given I_t:

Note that, since S_t and y_t are related by specification, knowing y_t can help to update $S_{t|t-1}$ by the following:

Write

$$\begin{aligned} S_t &= \hat{S}_{t|t-1} + (S_t - \hat{S}_{t|t-1}) \\ y_t &= A + B\hat{S}_{t|t-1} + B(S_t - \hat{S}_{t|t-1}) + v_t \end{aligned} \tag{40}$$

We have the following joint distribution:

$$\begin{bmatrix} S_t \\ y_t \end{bmatrix} |I_{t-1} \sim N(\begin{bmatrix} \hat{S}_{t|t-1} \\ A + B\hat{S}_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}B' \\ BP_{t|t-1} & BP_{t|t-1}B' + R \end{bmatrix}), \tag{41}$$

Thus,

$$\begin{aligned} \hat{S}_{t|t} &\equiv E[S_t|y_t, I_{t-1}] \\ &= \hat{S}_{t|t-1} + P_{t|t-1}B'(BP_{t|t-1}B' + R)^{-1}(y_t - BS_{t|t-1} - A) \end{aligned} \tag{42}$$

$$\begin{aligned} P_{t|t} &\equiv E[(S_t - \hat{S}_{t|t})(S_t - \hat{S}_{t|t})'] \\ &= P_{t|t-1} - P_{t|t-1}B'(BP_{t|t-1}B' + R)^{-1}BP_{t|t-1} \end{aligned} \tag{43}$$

4. Forecast S_{t+1} given I_t:

$$\begin{aligned} \hat{S}_{t+1|t} &= E[S_{t+1}|I_t] = F\hat{S}_{t|t} \qquad (44) \\ P_{t+1|t} &= E[(S_{t+1} - \hat{S}_{t+1|t})(S_{t+1} - \hat{S}_{t+1|t})'] \\ &= FP_{t|t}F' + Q_t \qquad (45) \end{aligned}$$

5. Maximum Likelihood Estimation of Parameters

The likelihood function can be built up recursively

$$\log L(Y_T) = \sum_{t=1}^{T} \log f(y_t|I_{t-1}), \qquad (46)$$

where

$$\begin{aligned} f(y_t|I_{t-1}) &= (2\pi)^{-1/2}|H'P_{t|t-1}H + R|^{-1/2} \\ &* \quad exp\{-\frac{1}{2}(y_t - A - B\hat{S}_{t|t-1})'(B'P_{t|t-1}B + R)^{-1}(y_t - A - B\hat{S}_{t|t-1})\} \\ for \quad t &= 1, 2, \ldots, T \qquad (47) \end{aligned}$$

Parameter estimates can then be based on the numerical maximization of the likelihood function.

References

1. Ait-Sahalia, Yacine (1995). "Testing Continuous-Time Models of the Spot Interest Rate," working paper, University of Chicago.

2. Backus, D. and S. Zin (1994). "Reverse Engineering the Yield Curve." unpublished paper, New York University, March.

3. Backus, D. and S. Zin (1993). "Long-Memory Inflation Uncertainty: Evidence from the Term Structure," *Journal of Money, Credit and Banking* 25, 3, part2 (August), 681-708.

4. Balduzzi, P., S.R. Das and S. Foresi (1995) "The Central Tendency: A Second Factor in Bond Yields", New York University, September 1995.

5. Bernanke, Ben S. and Alan S. Blinder (1992). "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review 82*, 901-921.

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

7. Brennan, M.J. and E.S. Schwartz (1982). "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency." *Journal of Financial and Quantitative Analysis* 17, 301-329.

8. Brown, R.H. and Schaefer (1991). "Interest Rate Volatility and the Term Structure." unpublished paper, London Business School.

9. Brown, S.J. and Dybvig (1986). "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates." *Journal of Finance.*

10. Brown, S.J. and M.I. Weinstein (1983). "A New Approach to Testing Asset Pricing Models: The Bilinear Paradigm" *Journal of Finance.*

11. Campbell, J.A. (1995). "Some Lessons from the Yield Curve." Journal of Economic Perspectives 9 (Summer), 129-152.

12. Campbell, J., A. Lo, and C. MacKinlay (1994). "Models of the Term Structure of Interest Rates." Federal Reserve Bank of Philadelphia Working Paper No. 94-10 (May).

13. Chan, K.C., A. Karolyi, F. Longstaff, and A. Sanders (1992). "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate." *Journal of Finance* 47, 1209-1227.

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

15. Duffie, Darrell (1992). Dynamic Asset Pricing Theory. Princeton, N.J.: Princeton University Press.

16. Eom, Y.H. (1995). "Essays on Empirical Tests of Continuous Time Asset Dynamnics." unpublished, New York University.

17. Engel, E.M.R.A. (1984). "A Unified Approach to the Study of Sums, Products, Time-Aggregation and Other Functions of ARMA Processes," *Journal of Time Series Analysis.* 5: 159-171.

18. Engsted, Tom (1995). "Does the Long-Term Interest Rate Predict Future Inflation? A Multi-Country Analysis," *Review of Economics and Statistics*, 42-54.

19. Estrella, A. and G. Hardouvelis (1991). "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46.

20. Estrella, A. and F.S. Mishkin (1995). "The Term Structure of Interest Rates and Its Role in Monetary Policy for the European Central Bank." *NBER* Working Paper 5279.

21. Fama, E.F.(1990). "Term-Structure Forcasts of Interest Rates, Inflation, and Real Returns." *Journal of Monetary Economics* 25, 59-76.

22. Fama, E.F. and R.R. Bliss (1987). "The Information in Long-Maturity Forward Rates." *American Economic Review* 77, 680-692.

23. Fisher, M., D. Nychka, and D. Zervos (1995). "Fitting the Term Structure of Interest Rates with Smoothing Splines." Finance and Economics Discussion Series, 95-1, Federal Reserve Board.

24. Gibbons, Michael R. and Krishna Ramaswamy (1993). "A Test of the Cox, Ingersoll, and Ross Model of the Term Structure." *The Review of Financial Studies* V.6, 619-658.

25. Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press.

26. Heath, David, Robert Jarrow, and Andrew Morton (1992). "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica* 60, 77-105.

27. Heston, S. (1992). "Testing Continuous-Time Models of the Term Structure of Interest Rates", unpublished paper, Yale University.

28. Ho, Thomas S.Y. and Sang-Bin Lee (1986). "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance* 41, 1011-1029.

29. Jegadeesh, N. and G.G. Pennacchi (1996). "The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures." Unpublished Paper, University of Illinois.

30. Litterman, R. and J. Scheinkman (1991). "Common Factors Affecting Bond Returns." *Journal of Fixed Income* (June) 54-61.

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

32. Lucas, Robert E. (1978). "Asset Prices in an Exchange Economy" *Econometrica* 46, No.6.

33. McCulloch, J. Huston and Heon-Chul Kwon (1993). "U.S. Term Structure Data, 1947-1991." Ohio State University Working Paper No. 93-6.

34. Mishkin, F. (1990). "What Does the Term Structure Tell Us About Future Inflation?" *Journal of Monetary Economics* 25, 77-95.

35. Sargent, Thomas J. (1987). Dynamic Macroeconomic Theory. Cambridge, Ma.: Harvard University Press.

36. Sims, Christopher A. (1996). "Macroeconomics and Methodology." *Journal of Economic Perspectives* Vol. 10, 105-120.

37. Singleton, Kenneth J. (1990). "Specification and Estimation of Intertemporal Asset Pricing Models," in Benjamin M. Friedman and Frank H. Hahn eds. Handbook of Monetary Economics. Amsterdam: North Holland.

38. Stanton, Richard (1996). "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk", unpublished paper, U.C., Berkeley.

39. 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.

40. Remolona, Eli M., Joseph Dziwura, and Irene Pedraza (1996). "Asymmetric Movements at the Short End of the Forward Curve." unpublished paper, Federal Reserve Bank of New York.

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

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