A THREE-FACTOR ECONOMETRIC MODEL OF THE U.S. TERM STRUCTURE

by
Frank F. Gong and Eli M. Remolona

**Federal Reserve Bank of New York
Research Paper No. 9619**

July 1996

This draft, July 1996
Not for quotation

A Three-Factor Econometric Model of the U.S. Term Structure

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 and test a model of the U.S. term structure that fits both the time series of interest rates and the cross-sectional shapes of the yield and volatility curves. In the model, three unobserved factors drive a stochastic discount process that prices assets so as to rule out arbitrage opportunities. The resulting bond yields are conveniently affine in the factors. We use monthly zero-coupon yield data from January 1986 to March 1996 and estimate the model by applying a Kalman filter that takes into account the model's no-arbitrage restrictions and using only three maturities at a time. The parameter estimates describe a first factor that reverts slowly to a fixed mean and a second factor that reverts relatively quickly to a time-varying mean serving as the third factor. The estimates are robust to the choice of maturities, suggesting that these factors give us an adequate model.

JEL Classification Codes: E43, G12, G13.

Keywords: Term structure, pricing kernel, affine yields, mean reversion, time-varying mean, Kalman filter.

[1]We thank Richard Cantor, John Clark, Arturo Estrella, Jean Helwege, Jim Mahoney, and Tony Rodrigues for helpful comments.

A Three-Factor Econometric Model of the U.S. Term Structure

FRANK F. GONG and ELI M. REMOLONA

Federal Reserve Bank of New York

I. Introduction

A challenge of equilibrium models of the term structure of interest rates is to reconcile the time-series dynamics of interest rates with the cross-sectional shapes of the yield and volatility curves. Backus and Zin (1992) and Campbell, Lo, and MacKinlay (1994, hereafter CLM) have pointed out that such models have been estimated to be consistent with either the time series or the cross section of bond yields but not both.[2] When the models are based on time-series data, the constructed term structures fail to match the shapes of actual curves.[3] When the models are based on cross-section data, parameter values must vary over time with shifts in the term structure.[4] In this paper, we specify and estimate an equilibrium model of the U.S. term structure that is consistent with both time-series and cross-section data. The model we propose

[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]To market participants, it is more critical that their model be consistent with the cross section than with the time series, particularly for pricing contingent claims. However, as Black and Karasinski (1991) point out, relying solely on cross section data means having a different model from one moment to the next. The most popular cross-section models are Ho and Lee (1986), Black, Derman, and Toy (1990), Hull and White (1990), and Heath, Jarrow, and Morton (1992).

is one that requires three factors to price bonds consistently across the term structure and produce the actual shapes of yield and volatility curves.

The number of factors required for an adequate term structure model is an important issue. The point is to build a consistent model with as few factors as possible. Litterman and Scheinkman (1991) show that three factors can explain nearly all the variation in bond returns. They interpret their factors as representing the level of interest rates, the slope of the yield curve, and the curvature of the yield curve. However, they do not ensure that their factor loadings are consistent with no arbitrage. Gong and Remolona (1996) are careful to impose no arbitrage when they fit three alternative two-factor models to U.S. quarterly yield data. However, they fail to find a model that is adequate for explaining the whole term structure, and they conclude that at least three factors would be required for that purpose. Chen (1996) proposes but does not estimate a three-factor model in which the future short-term interest rate is determined by its current value, its time-varying mean, and its stochastic volatility.

We follow Backus and Zin (1992) and CLM (1994) by specifying a term structure model in terms of a stochastic discount process. We use this process, known as a pricing kernel, to consistently price bonds of different maturities so that we avoid arbitrage opportunities. To maintain tractability, we write the model to satisfy Duffie and Kan's (1993) conditions for affine yields. In this model, three unobserved factors drive the pricing kernel: one factor reverts over time to a fixed mean while a second factor reverts to a time-varying mean that serves as the third factor. Such a model will produce reduced forms that nest most factor models in the literature.[5] To capture the hump in the yield

[5]Examples of these models are the one-factor models of Vasicek (1977) and Cox, Ingersoll,

curve, we follow Gong and Remolona (1996) by pricing a risk associated with the volatility of the time-varying mean.

To estimate and test the model, we use monthly U.S. Treasury zero-coupon yield data from January 1986 to March 1996. The model lends itself to estimation by a Kalman filter, and we apply the technique in a way that takes account of the model's arbitrage conditions. The yields as functions of the factors serve as the measurement equations and the factors' stochastic processes as the transition equations. The arbitrage conditions impose restrictions between the measurement and transition equations. To test the adequacy of three factors, we estimate the model using only three yields at a time – a short-term rate, a medium-term rate, and a long-term rate. The estimates turn out to be robust to the choice of maturities, suggesting that three factors are adequate.

The fit between the estimated model and actual yield and volatility curves strikes us as impressive. In particular, the unconditional yield and volatility curves we construct from our parameter estimates capture the hump in the average yield curve and the flatness of the average volatility curve for the sample period. The shapes of these curves are most sensitive to the rates of mean reversion in the factors, the price of risk, and the volatility of the time-varying mean. Our estimates describe a first factor that reverts rather slowly to a fixed mean and a second factor that reverts relatively quickly to a time-varying mean. The first factor's slow mean reversion decides the yield curve's slope near the long end and the volatility curve's slope across most of its length. The second factor's fast mean reversion determines the yield curve's

and Ross (1985) and the two-factor models of Brennan and Schwartz (1979), Schaefer and Schwartz (1984), Longstaff and Schwartz (1992), and Campbell, Lo, and Mackinlay (1994).

slope near the short end. The price of risk and the third factor's volatility impart curvature to the yield curve. We estimate a relatively high price of risk (in absolute value) and a relatively low volatility, and these estimates produce just enough curvature to capture the hump in the actual yield curve.

In what follows, we begin with a brief discussion of pricing-kernel affine-yield models of the term structure. In Section III, we then illustrate with monthly data the inadequacy of two-factor affine yield models, as Gong and Remolona have demonstrated with quarterly data. We specify our three-factor model in Section IV, estimate it in Section V, and evaluate the robustness of the estimates in Section VI. We discuss the role of the various parameters in shaping the yield and volatility curves in Section VII. We propose further work in section VIII.

II. Theory: Affine Yield Pricing Kernel Models

A. Background Literature

Theoretical work with equilibrium models, notably by Vasicek (1977) and Cox, Ingersoll, and Ross (1985, hereafter CIR), show how the term structure at a moment in time would reflect regularities in interest rate movements over time. In the simplest such models, the short-term interest rate is the single factor driving movements in the term structure. Vasicek assumes that the short-rate's volatility is constant, while CIR assume that it is proportional to the square root of the short rate itself. The absence of arbitrage requires that the ratio of expected excess return to return volatility be the same for different bonds. This arbitrage condition, the assumption of lognormal bond prices, and either Vasicek's or CIR's short-rate volatility produce an affine-yield solution in which all bond yields (or log bond prices) are linear functions of the short-term rate. Such linearity simplifies the pricing of fixed-income securities and contingent claims. In the two-factor models of Brennan and Schwartz (1979), Schaefer and Schwartz (1984), Longstaff and Schwartz (1992), and CLM (1994) and in Chen's (1996) three-factor model, similar assumptions generate bond yields that are also linear in the factors. Duffie and Kan (1993) establish the conditions that produce such affine yields in general.

Rather than model the short-term interest rate directly, Backus and Zin (1992) and CLM (1994) focus on the stochastic discount process or the pricing kernel used to price assets in general. Arbitrage opportunities are avoided by applying the same pricing kernel to different assets. In this approach, the factors are unobservable state variables that serve to forecast discount rates. In principle, the factors can be related to observable macroeconomic fundamentals, as Gong and Remolona (1996) try to do. Pricing kernel models can also be specified so that bond yields are affine in the factors and, with a linear transformation, affine in the short rate as well. We describe below such a pricing-kernel affine-yield model with K factors.

B. The Pricing Kernel

The pricing kernel approach relies on a no-arbitrage condition common to intertemporal asset pricing models.[6] In the case of zero-coupon bonds, the price of an n-period bond is

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

[6] Singleton (1990) provides a critical survey of these models, particularly their empirical performance. Duffie (1992) relates arbitrage conditions to concepts of optimality and equilibrium.

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. 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.[7] 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.

C. K-Factor Affine Yields

For our affine yield models, we assume that 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}.

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 K factors, $x_{1t}, x_{2t}, \ldots, x_{Kt}$, that forecast m_{t+1}, an affine yield model can be written as

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

Since the n-period bond yield is $y_{nt} = -p_{nt}/n$, yields will also be linear in the factors. The coeficients $A_n, B_{1n}, B_{2n}, \ldots, B_{Kn}$ will depend on the stochastic

[7]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).

processes of $x_{1t}, x_{2t}, \ldots, x_{Kt}$. Since the number of factors is usually smaller than the number of maturities on the curve, the factor structure would imply restrictions across coefficients for bond prices of different maturities. In practice, specifying $A_n, B_{1n}, B_{2n}, \ldots, B_{Kn}$ involves solving (2) based on the stochastic processes of $x_{1t}, x_{2t}, \ldots, x_{Kt}$ and verifying that (3) holds.

III. Failure of Two-Factor Affine Models

A. Previous results

In the effort to reconcile time-series data with cross-section data on interest rates, models with fewer than three factors have not fared well. Backus and Zin (1992) and CLM (1994) argue that the basic problem with one-factor models is that the yield curve's steep slope near the short end requires swift mean reversion by the factor while the curve's flat slope near the long end requires slow mean reversion. The flat slope of the volatility curve also requires slow mean reversion. Gong and Remolona (1996) demonstrate that the problem is not solved with two factors either. They find that the data favor models in which one of the factors is a time-varying mean, which serves to produce the characteristic hump in the U.S. yield curve around the one-year to two-year maturities. The other factor must then revert rapidly to this mean to create a steep yield curve near the short end but revert slowly to create both a flat yield curve near the long end and a flat volatility curve.

In this section, we replicate Gong and Remolona's results with a somewhat different data set. We use monthly data on U.S. zero-coupon Treasury yields from January 1986 to March 1996. Gong and Remolona used quarterly data on those yields from 1984 Q1 to 1995 Q4. As they did, we try to fit a two-factor additive model and a two-factor time-varying mean model using two yield maturities at a time and using a Kalman filter that takes into account the models' no-arbitrage restrictions. If two factors are adequate, then we should be able to estimate the same model with any two maturities. We discuss the estimation procedure in more detail in Section V, where we turn to the estimation of a three-factor model.

B. A two-factor additive model

In the additive model, the conditional expectation of the negative of the log stochastic discount factor depends on two factors that enter additively:

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

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}. 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 can be solved to produce affine yields

$$y_t^n = \frac{1}{n}(A_n + B_{1n}x_{1t} + B_{2n}x_{2t}) \tag{7}$$

where A_n, B_{1n}, and B_{2n} depend on $\phi_1, \phi_2, \mu, \theta, \lambda, \sigma_1^2, \sigma_2^2$, and σ_{12}.

C. A time-varying mean model

In the 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:

$$\begin{aligned} -m_{t+1} &= x_t + w_{t+1} \\ x_{t+1} &= (1-\phi_1)\mu_t + \phi_1 x_t + \mu_t^{0.5} u_{1,t+1} \\ \mu_{t+1} &= (1-\phi_2)\theta + \phi_2 \mu_t + \mu_t^{0.5} u_{2,t+1}, \\ w_{t+1} &= \lambda \mu_t^{0.5} u_{2,t+1} \end{aligned} \tag{8}$$

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 solution to the model produces affine yields

$$y_t^n = \frac{1}{n}(A_n + B_{1n}x_{1t} + B_{2n}\mu_t) \tag{9}$$

where A_n, B_{1n}, and B_{2n} depend on $\phi_1, \phi_2, \theta, \lambda, \sigma_1^2, \sigma_2^2$, and σ_{12}.

D. *The poor fit of two-factor models*

With only two factors, the constructed yield and volatility curves tend to match the actual curves only at the maturities used in estimation. Moreover, the estimates are not robust to the choice of maturities. Figs. 1 to 6 compare the curves implied by the estimated models to the actual averaage curves for the sample period. In Figs. 1 and 2, the yield and volatility curves implied by the models are based on estimates using only three-month and two-year yields. The models produce yields that fall below the actual yields for maturities longer than two years, with the additive model even producing negative yields for maturities longer than eight years. At the long end, the volatilities implied by the additive model are too high and those implied by the time-varying mean model too low. In Fig. 3, the models are estimated on two-year and ten-year yields, and the implied yields are too low near the short end and too high in the maturities between two and ten years. In Fig. 5, the models are estimated on three-month and ten-year yields, and the implied yields are too high in the intermediate maturities. The evidence suggests that two factors will not give us an adequate model.

IV. A Three-factor Model

We now propose a three-factor model to fit both the time-series dynamics of interest rates and the cross-sectional shapes of the term structure. We follow Backus and Zin (1992), CLM (1994), and Gong and Remolona (1996) by specifying the model in terms of a pricing kernel. To maintain tractability, we write the model to satisfy Duffie and Kan's (1993) conditions for affine yields. In this model, three unobserved factors drive the pricing kernel: one factor reverts over time to a fixed mean while a second factor reverts to a time-varying mean that serves as the third factor. The model is a combination of CLM's two-factor additive model and Gong and Remolona's time-varying mean model. From the outset, we specify the model in discrete time to avoid possible problems in estimating a continuous-time model with discrete-time data.

A. Model specification

Three unobservable factors drive the pricing kernel. Two of the factors directly affect expectations of the stochastic discount factor for the next period, while the third factor affects the ultimate destination of the stochastic discount factor. Specifically, the conditional expectation of the negative of the log stochastic discount factor depends on the sum of two factors:

$$-m_{t+1} = x_{1t} + x_{2t} + w_{t+1}, \tag{10}$$

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 time-varing mean of the second factor. Each of these factors follows a univariate AR(1) process with heteroscedastic shocks described by a square-root process:

$$\begin{aligned} x_{1,t+1} &= (1-\phi_1)\theta + \phi_1 x_{1,t} + \mu_t^{0.5} u_{1,t+1} \\ x_{2,t+1} &= (1-\phi_2)\mu_t + \phi_2 x_{2,t} + \mu_t^{0.5} u_{2,t+1} \\ \mu_{t+1} &= (1-\phi_3)\mu + \phi_3 \mu_t + \mu_t^{0.5} u_{3,t+1}, \end{aligned} \tag{11}$$

where $1-\phi_1$, $1-\phi_2$, and $1-\phi_3$ are the rates of mean reversion, θ and μ are the long-run means to which the factors revert, and $u_{1,t+1}$, $u_{2,t+1}$, and $u_{3,t+1}$ are shocks with mean zero, volatilities σ_1, σ_2 and σ_3 and covariances $\sigma_{12}, \sigma_{13}, \sigma_{23}$.

It is important to allow correlation between factor shocks, if we hope to relate the factors to fundamentals, which may not be orthogonal.

As in Gong and Remolona (1996), we specify the shock to m_{t+1} to be proportional to the shock to μ_{t+1}, which in turn depends on the level of μ_t:

$$w_{t+1} = \lambda \mu_t^{0.5} u_{3,t+1} \tag{12}$$

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.

We now verify that yields are affine in the factors so that we can write

$$y_t^n = \frac{1}{n}(A_n + B_{1n}x_{1t} + B_{2n}x_{2t} + B_{3n}\mu_t). \tag{13}$$

The normalization $p_{0t} = 0$ gives us coefficients of $A_0 = B_{10} = B_{20} = B_{30} = 0$. We can then derive the one-period yield or short rate as

$$\begin{aligned} y_{1,t} &= -p_{1,t} = -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\ &= x_{1,t} + x_{2,t} - \frac{1}{2}\lambda^2\sigma_3^2\mu_t \end{aligned} \tag{14}$$

which is also linear in the factors, with the coefficients $A_1 = 0$, $B_{11} = B_{21} = 1$, and $B_{31} = -\frac{1}{2}\lambda^2\sigma_3^2$.

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

$$\begin{aligned} A_n &= A_{n-1} + (1-\phi_1)\theta B_{1,n-1} + (1-\phi_3)\mu B_{3,n-1} \\ B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\ B_{2,n} &= 1 + \phi_2 B_{2,n-1} \\ B_{3,n} &= \phi_3 B_{3,n-1} + (1-\phi_2)B_{2,n-1} \end{aligned} \tag{15}$$

$$
\begin{aligned}
&- \frac{1}{2}[(\lambda + B_{3,n-1})^2\sigma_3^2 + B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 \\
&+ 2(\lambda + B_{3,n-1})B_{1,n-1}\sigma_{13} + 2(\lambda + B_{3,n-1})B_{2,n-1}\sigma_{23} + 2B_{1,n-1}B_{2,n-1}\sigma_{12}],
\end{aligned}
$$

The coefficients B_{1n}, B_{2n}, and B_{3n} are factor loadings for x_{1t}, x_{2t}, and μ_t. 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 twelve parameters: $\phi_1, \phi_2, \phi_3, \theta, \mu, \sigma_1, \sigma_2, \sigma_3, \sigma_{12}, \sigma_{13}, \sigma_{23}$, and λ.

To price fixed-income options, we need a consistent volatility curve. In the case of our model, such a curve is derived from the conditional variance of the n-period yield:

$$
\begin{aligned}
Var_t(y_{n,t+1}) &= \frac{1}{n^2}(B_{1n}^2\sigma_1^2 + B_{2n}^2\sigma_2^2 + B_{3n}^2\sigma_3^2 \\
&+ 2B_{1n}B_{2n}\sigma_{12} + 2B_{1n}B_{3n}\sigma_{13} + 2B_{2n}B_{3n}\sigma_{23})\mu_t \qquad (16)
\end{aligned}
$$

We would have a downward-sloping volatility curve given that ϕ_1, ϕ_2, and ϕ_3 are less than unity. Mean reversion by the factors serves to dampen yield volatilities as maturity is lengthened.

A linear transformation of this model will give us a reduced form which which expresses the yield for a given maturity as a linear function of some other yield, the conditional variance of some yield, and the conditional expectation of some yield. This follows from the fact that yields, conditional variances, and conditional expectations are different linear functions of the factors. Such a reduced form will nest the one-factor models of Vasicek (1977) and CIR (1985) and the two-factor models of Brennan and Schwartz (1979), Brennan and Schaefer (1984), and Longstaff and Schwartz (1994). One such reduced form will also look like Chen's (1996) three-factor model.

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

$$\begin{aligned}
E_t p_{n-1,t+1} - p_{nt} \; - \; & y_{1t} = -\lambda(B_{1,n-1}\sigma_{13} + B_{2,n-1}\sigma_{23} + B_{3,n-1}\sigma_3^2)\mu_t \\
- \; & \frac{1}{2}[B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + B_{3,n-1}^2\sigma_3^2 \qquad (17) \\
+ \; & 2B_{1,n-1}B_{2,n-1}\sigma_{12} + 2B_{1,n-1}B_{3,n-1}\sigma_{13} + 2B_{2,n-1}B_{3,n-1}\sigma_{23}]\mu_t \\
= \; & -(\lambda\Pi + \Gamma)\mu_t
\end{aligned}$$

where

$$\Pi \equiv (B_{1,n-1}\sigma_{13} + B_{2,n-1}\sigma_{23} + B_{3,n-1}\sigma_3^2) \qquad (18)$$

$$\begin{aligned}
\Gamma \equiv \; & \frac{1}{2}(B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + B_{3,n-1}^2\sigma_3^2 \\
+ \; & 2B_{1,n-1}B_{2,n-1}\sigma_{12} + 2B_{1,n-1}B_{3,n-1}\sigma_{13} + 2B_{2,n-1}B_{3,n-1}\sigma_{23})\mu_t \qquad (19)
\end{aligned}$$

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

$$-\lambda > \frac{\Gamma}{\Pi} \qquad (20)$$

Note also that if $\sigma_3 = 0$, we will have homoscedastic shocks, term premia will be constant, and the pure expectations hypothesis will hold.

V. Estimating the Model

A. Econometric Approach

Our econometric approach allows us to estimate the parameters of a pricing kernel without directly observing it or the three factors that are supposed to drive it. Moreover, to confront the issue of model adequacy, we estimate the model using yields of only three maturities at a time. If a three-factor model is to explain the movements of the whole term structure, then we should be able to estimate it with only three maturities, and the model should be robust to the choice of maturities. Our approach is a special application of the Kalman filter. In applying this technique, we make use of observed yields, which would reflect the dynamics of the factors, and we use only three yields at a time – a short-term yield, a medium-term yield, and a long-term yield. The yields as affine functions of the factors serve as the measurement equations of the Kalman filter and the factors' stochastic processes as the transition equations. The model's arbitrage conditions, however, imply strong restrictions between the measurement and transition equations, and we take careful account of these restrictions.

Recent efforts to reconcile time series data with cross section data on interest rate have tended to rely on the GMM approach. In practice this approach places cross-section restrictions on only some of the unconditional 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 when they fit and test the CIR model. 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 six parameters implied by their structural model.

For our purposes, the Kalman filter is the appropriate estimation procedure. The technique is effective in exploiting conditional moments, which constitute essential information when one is trying to estimate the dynamics of three unobserved factors on the basis of observed yields for only three maturities. The technique is especially suitable for estimating term structure models, because it allows the imposition of the arbitrage restrictions. Jegadeesh and Pennacchi (1996) use the Kalman filter in estimating a two-factor term structure model using data on Eurodollar futures, although they use more than two yields at a time. Gong and Remolona (1996) use the technique in their exploration of the U.S. term structure with two-factor models, and they use only two yields at a time.

The work builds on Gong and Remolona, and in spirit, it 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 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 $ARMA(3,2)$ yield process is generated by our three $AR(1)$ factors.[8]

B. Data and Summary Statistics

We obtained end-of-month U.S. zero-coupon Treasury yield data for maturities of one-year and longer from J.P. Morgan and Company and for maturities of three months and six months from the Federal Reserve Bank of New York. The sample period is 1986:1 to 1996:3.[9] 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

[8]Engel (1984) derives the sums, products, and time aggregations of ARMA processes.

[9]The data are available on request.

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 3 months, 6 months, 1 year, 2, 5, and 10 years for the sample period 86:1-96:3 are reported in Table 1. The average term structure is upward sloping, with mean yields ranging from 5.54% to 7.80%. Its slope is steep near the short end and flat near the long end. This term structure is some what hump-shaped, with the hump located near the two year maturity. Overall, the volatility curve slopes downward very gradually. The yields across the curve are all very persistent, with first-order monthly auto-correlations of 0.94-0.99.

To evaluate robustness, we use four different combinations of yield maturities to estimate the parameters. The four different combinations are: 3-month, 2-year, and 10-year yields; 3-month, 1-year, and 10-year yields; 6-month, 2-year, and 10-year yields; and 3-month, 2-year, and 5-year yields. If the three-factor model is adequate for explaining the whole term-structure, we should get similar parameter estimates and implied yield and volatility curves from the different combinations of maturities.

C. Kalman Filtering and Maximum Likelihood Estimation

We now show how to fit the three-factor model to U.S. zero-coupon rates data, using three-yields at a time.

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

$$\begin{bmatrix} y_{l,t} \\ y_{m,t} \\ y_{n,t} \end{bmatrix} = \begin{bmatrix} a_l \\ a_m \\ a_n \end{bmatrix} + \begin{bmatrix} b_{1,l} & b_{2,l} & b_{3,l} \\ b_{1,m} & b_{2,m} & b_{3,m} \\ b_{1,n} & b_{2,n} & b_{3,n} \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \mu_t \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{bmatrix} \tag{21}$$

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

$$v_t \sim N(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} e_1^2 & 0 & 0 \\ 0 & e_2^2 & 0 \\ 0 & 0 & e_3^2 \end{pmatrix}), \tag{22}$$

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

The transition equation is

$$\begin{bmatrix} x_{1,t+1} \\ x_{2,t+1} \\ \mu_{t+1} \end{bmatrix} = \begin{bmatrix} (1-\phi_1)\theta \\ 0 \\ (1-\phi_3)\mu \end{bmatrix} + \begin{bmatrix} \phi_1 & 0 & 0 \\ 0 & \phi_2 & 1-\phi_2 \\ 0 & 0 & \phi_3 \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \mu_t \end{bmatrix} + \mu_t^{0.5} \begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \\ u_{3,t+1} \end{bmatrix}$$

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

$$\begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \\ u_{3,t+1} \end{bmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23}^2 & \sigma_3^3 \end{pmatrix}) \tag{23}$$

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 (15).

After putting the restrictions into the measurement equations, the preceding models can be estimated by maximum likelihood using the Kalman filter.

The algorithm is discussed in Appendix C. For more detailed discussions of the Kalman filtering procedure, see, for example, Hamilton (1994).

VI. Estimates of the Model and Robustness

Our estimates allow us to test the adequacy of our model and to differentiate among the three factors that are supposed to be driving the model. We informally test the model's adequacy by comparing alternative parameter estimates based on different combinations of yield maturities and by examining how well the unconditional yield and volatility curves produced by the estimates fit the actual average curves for the sample period. We differentiate among the factors primarily by comparing their estimated rates of mean reversion. Table 3 reports our parameter estimates based on the four alternative combinations of yield maturities. It is immediately apparent that the different combinations of maturities provide very similar parameter estimates. These estimates characterize a first factor that reverts slowly to fixed mean, a second factor that reverts rapidly to the third factor, which itself reverts more rapidly to its own mean than does the first factor and more slowly than does the second factor.

Since we estimate the model with the observed yields for only three maturities at a time, we may evaluate the model by comparing estimates based on other maturities. If a three-factor model held, then the estimates should be robust to the choice of maturities. Table 3 compares estimates based on the following maturity combinations: (1) the three-month, two-year, and ten-year yields; (2) the three-month, one-year, and ten-year yields; (3) the six-month, two-year, and ten-year yields; and (4) the three-month, two-year, and five-year yields. The differences among the alternative parameter estimates are statistically insignificant. The similarity among the alternative estimates is particularly impressive for the parameters that are tightly estimated, such as the persistence parameters ϕ_1, ϕ_2, and ϕ_3. The parameter estimates are never more then two standard errors apart even based on the smallest standard error of 0.006 for the estimate of ϕ_1 using the three-month, two-year, and ten-year yields.

Table 4 reports estimates based on two subsamples, one for January 1986 to December 1980 and one for January 1991 to March 1996. Again the parameter estimates are remarkably similar across subsamples.

We may also evaluate the model by seeing how well it reproduces the rest of the term structure. Our Kalman filter procedure would tend to cause the term structure to match at the maturities used to estimate the model. However, the procedure will not assure a fit with the rest of the term structure unless the yields for other maturities reflect the dynamics of the same three factors. Fig. 7 compares the unconditional yield curves we construct from our alternative model estimates with the actual average yield curve for the sample period. Similarly, Fig. 8 compares the unconditional volatility curves from the alternative estimates with the actual average volatility curve. On the whole, the fit between the estimated model and actual yield and volatility curves is remarkable, especially compared to the results of the two factor models. In particular, the implied yield and volatility curves capture the hump in the average yield curve and the flatness of the average volatility curve for the sample period.

How do the factors differ? Recall that the rate of mean reversion is given by $1 - \phi_1$ for the first factor, $1 - \phi_2$ for the second factor, and $1 - \phi_3$ for the third factor. As Table 3 reports, the first factor reverts to its fixed mean at the rather slow rate of one percent a month. This rate of mean reversion implies a near unit root process.[10] The second factor reverts to its time-varying mean at the rate of about 12 percent a month or a rate 12 times faster than that of the first factor. The third factor, which is the time-varying mean, reverts to its own mean at the rate of five percent, which is five times faster than the first factor but less than half as fast as the second factor. Such mean reversion rates are critical determinants of the shape of the term structure. The volatility estimates suggest that the first factor has the smallest shocks while the second factor has the largest ones. Shocks to the first factor are not significantly correlated with shocks to the second or third factor. The second and third factors, however, have shocks that are significantly positively correlated.

VII. How the Factors Shape the Term Structure

How do the three factors succeed in reproducing the actual shapes of the term structure? The average U.S. yield curve can be characterized as having a steep slope near the short end, a flat slope near the long end, and something like a hump around the one-year to two-year maturities. The average U.S. volatility curve can be characterized as downward sloping but with a rather flat slope. Our estimates of the three-factor model suggest that the first factor accounts for yield curve's flat slope near the long end and the volatility curve's flat slope for most of its length. The second factor accounts for the yield curve's

[10]We can rule out a unit root process, however, because such a process implies negative yields at long maturities. If $\phi_1 = 1$, we will have $B_{1n} = n$, and the term premium in (17)-(19) must eventually turn negative.

steep slope near the short end. The third factor combined with the price of risk impart just enough curvature to account for the yield curve's hump.

The slopes of the yield curve near either end are the easiest characteristics of the term structure to explain. The first factor determines the slope near the long end, because it is the factor with the slowest mean reversion. It would have the most influence on the long end because a shock to the factor would persist the longest and would be the shock most likely to be reflected in long-term yields. The second factor determines the slope near the short end, because it is the factor with the most rapid mean reversion, and it would have its greatest effect on the shorter term yields. These explanations are consistent with the factor loadings shown in Fig. 9. The picture portrays the first factor as having an effect on yields that decays very slowly as maturity is lengthened and the second factor as having an effect that dies down so swiftly that its effect at the ten-year maturity is only a fifth of that of the first factor.

The hump in the yield curve is apparently a feature associated with the time-varying mean and the effect it has on the risk premium. Such an association is evident in time-varying mean models with two factors as well as those with three factors. In these models, the time-varying mean factor induces hetorescedasticity in volatility, and this source of risk is priced. In the present three-factor model, the loading for this factor, as shown in Fig. 9, is initially negative then becomes positive around the 16-month maturity, which is roughly the location of yield curve's hump. To produce the right curvature, however, requires the right mean reversion rate as well as the right volatility and price of risk. Fig. 10, for example, shows the implied yield curve produced with the factor having a mean reversion rate of one percent instead of five percent. The curve overshoots the average one-year yield and undershoots the yields between the two-year and ten-year maturities.

In general, a downward sloping volatility curve requires mean reversion and a flat curve slow mean reversion. To produce the right shape for this curve, however, requires slow mean reversion for the first factor only. Fig. 11, for example, shows that a slow rate of mean reversion for the time-varying mean induces a volatility curve that overshoots volatilities for maturities between two years and ten years.

VIII. Conclusion

We believe we have an adequate econometric model of the U.S. term structure. It is a model of a pricing kernel that serves to consistently price bonds of different maturities so that arbitrage opportunities do not arise. Three factors drive this pricing kernel: one factor reverts over time to a fixed mean, a second factor reverts to a time-varying mean, and the time-varying mean itself is a mean-reverting factor that induces time-varying term premia. With alternative estimates of the model using a special Kalman filter and only three maturities at a time, we find that different combinations of maturities produce the same three factors, particularly as characterized by their rates of mean reversion. The estimates also reproduce the actual average yield and volatility curves for the sample period, suggesting that the yields not used in estimation also reflect the time-series dynamics of the same three factors.

The estimates describe three factors with very different rates of mean reversion. The first factor reverts to its fixed mean at the rate of about one percent a month, the second factor reverts to its time-varying mean about 12 times faster; and the third factor reverts to its own mean five times faster than the first factor but less than half as fast as the second factor. Something seems key about these parameter values, because small deviations from our range of estimates produce very bad yield and volatility curves. Each factor has a role in the shapes of the term structure. The first factor explains the yield curve's flat slope near the long end and the volatility curve's flat slope for most of its length. The second factor explains the yield curve's steep slope near the short end. The time-varying mean produces the yield curve's hump.

There is more work to be done. Having decided that three factors are adequate, we would like to estimate the model using all the maturities available at once. Such an estimate will be more efficient than the ones reported in this paper and will be appropriate for testing the significance of the arbitrage restrictions. We would also like to use such a model to forecast changes in short-term rates and long-term rates over different time horizons to see whether controlling for time-varying term premia by means of the model would support the expectations hypothesis.

Appendix A

A1. Model I: Recursive Restrictions

We start with the general pricing equation:

$$p_{nt} = E_t(m_{t+1} + p_{n-1,t+1}) + \frac{1}{2}Var_t(m_{t+1} + p_{n-1,t+1}) \tag{24}$$

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

$$\begin{aligned} y_{1t} = -p_{1t} &= -E_t(m_{t+1}) - \frac{1}{2}Var_t(m_{t+1}) \\ &= x_{1,t} + x_{2,t} - \frac{1}{2}\lambda^2\sigma_3^2\mu_t, \end{aligned}$$

showing the short rate to be linear in the factors.

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

$$-p_{n,t} = A_n + B_{1,n}x_{1,t} + B_{2,n}x_{2,t} + B_{3,n}\mu_t \tag{25}$$

We verify that there exist A_n, $B_{1,n}$, $B_{2,n}$, and $B_{3,n}$ that satisfy the general pricing equation:

$$\begin{aligned} E_t(m_{t+1} + p_{n-1,t+1}) &= -A_{n-1} - (1-\phi_1)\theta B_{1,n-1} - (1-\phi_3)\mu B_{3,n-1} \\ &- (1+\phi_1 B_{1,n-1})x_{1,t} - (1+\phi_2 B_{2,n-1})x_{2,t} \\ &- [(1-\phi_2)B_{2,n-1} + \phi_3 B_{3,n-1}]\mu_t \qquad (26) \\ Var_t(m_{t+1} + p_{n-1,t+1}) &= [(\lambda + B_{3,n-1})^2\sigma_3^2 + B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 \\ &+ 2(\lambda + B_{3,n-1})B_{1,n-1}\sigma_{13} + 2B_{1,n-1}B_{2,n-1}\sigma_{12} \\ &+ 2(\lambda + B_{3,n-1})B_{2,n-1}\sigma_{23}]\mu_t \qquad (27) \end{aligned}$$

Now substitute (26) (27) into (24) and match coefficients of equations (24) and (25), we have

$$A_n = A_{n-1} + (1-\phi_1)\theta B_{1,n-1} + (1-\phi_3)\mu B_{3,n-1}$$

$$
\begin{aligned}
B_{1,n} &= 1 + \phi_1 B_{1,n-1} \\
B_{2,n} &= 1 + \phi_2 B_{2,n-1} \\
B_{3,n} &= \phi_3 B_{3,n-1} + (1-\phi_2) B_{2,n-1} \\
&- \frac{1}{2}[(\lambda + B_{3,n-1})^2 \sigma_3^2 + B_{1,n-1}^2 \sigma_1^2 + B_{2,n-1}^2 \sigma_2^2 \\
&+ 2(\lambda + B_{3,n-1}) B_{1,n-1} \sigma_{13} + 2(\lambda + B_{3,n-1}) B_{2,n-1} \sigma_{23} + 2 B_{1,n-1} B_{2,n-1} \sigma_{12}],
\end{aligned}
$$

Appendix B

B1. The Term Premia

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

$$
\begin{aligned}
E_t & \quad p_{n-1,t+1} - p_{n,t} - y_{1t} \\
= & \quad -A_{n-1} - B_{1,n-1}E_t x_{1,t+1} - B_{2,n-1}E_t x2,t+1 - B_{3,n-1}E_t\mu_{t+1} \\
+ & \quad A_n + B_{1,n}x_{1,t} + B_{2,n}x_{2,t} + B_{3,n}\mu_t \\
- & \quad x_{1,t} - x_{2,t} + \frac{1}{2}\lambda^2\sigma_3^2\mu_t \\
= & \quad (A_n - A_{n-1} - B_{1,n-1}(1-\phi_1)\theta - B_{3,n-1}\mu(1-\phi_3)\mu) \\
+ & \quad (B_{1,n} - 1 - \phi_1 B_{1,n-1})x_{1t} + (B_{2,n} - 1 - \phi_2 B_{2,n-1})x_{2t} \\
+ & \quad (B_{3n} + \frac{1}{2}\lambda^2\sigma_3^2 - \phi_3 B_{3,n-1} - (1-\phi_2)B_{2,n-1})\mu_t \\
= & \quad -\lambda(B_{1,n-1}\sigma_{13} + B_{2,n-1}\sigma_{23} + B_{3,n-1}\sigma_3^2)\mu_t \\
- & \quad \frac{1}{2}[B_{1,n-1}^2\sigma_1^2 + B_{2,n-1}^2\sigma_2^2 + B_{3,n-1}^2\sigma_3^2 \\
+ & \quad 2B_{1,n-1}B_{2,n-1}\sigma_{12} + 2B_{1,n-1}B_{3,n-1}\sigma_{13} + 2B_{2,n-1}B_{3,n-1}\sigma_{23}]\mu_t
\end{aligned}
\tag{28}
$$

B2. Expected Change in the Short Rate

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

$$
E_t y_{1,t+n} = E_t x_{1,t+n} + E_t x_{2,t+n} - \frac{1}{2}\lambda^2\sigma_3^2 E_t\mu_{t+n}
$$

$$
E_t \begin{bmatrix} x_{1,t+1} - \theta \\ x_{2,t+1} - \mu \\ \mu_{t+1} - \mu \end{bmatrix} = \begin{bmatrix} \phi_1 & 0 & 0 \\ 0 & \phi_2 & 1-\phi_2 \\ 0 & 0 & \phi_3 \end{bmatrix} \begin{bmatrix} x_{1,t} - \theta \\ x_{2,t} - \mu \\ \mu_t - \mu \end{bmatrix}
$$

$$
\begin{aligned}
E_t \begin{bmatrix} x_{1,t+n} - \theta \\ x_{2,t+n} - \mu \\ \mu_{t+n} - \mu \end{bmatrix} &= \begin{bmatrix} \phi_1 & 0 & 0 \\ 0 & \phi_2 & 1-\phi_2 \\ 0 & 0 & \phi_3 \end{bmatrix}^n \begin{bmatrix} x_{1,t} - \theta \\ x_{2,t} - \mu \\ \mu_t - \mu \end{bmatrix} \\
&= \begin{bmatrix} \phi_1^n & 0 & 0 \\ 0 & \phi_2^n & \frac{1-\phi_2}{\phi_2-\phi_3}(\phi_2^{n+1} - \phi_3^{n+1}) \\ 0 & 0 & \phi_3^n \end{bmatrix} \begin{bmatrix} x_{1,t} - \theta \\ x_{2,t} - \mu \\ \mu_t - \mu \end{bmatrix}
\end{aligned}
$$

We can then write

$$
E_t y_{1,t+n} - y_{1,t} = \alpha_n + \beta_{1n}(\theta - x_{1t}) + \beta_{2n}(\mu_t - x_{2t}) + \gamma_n \mu_t
$$

where

$$
\begin{aligned}
\alpha_n &= [1 - \phi_2^n - \frac{1-\phi_2}{\phi_2 - \phi_3}(\phi_2^{n+1} - \phi_3^{n+1}) - \frac{1}{2}\lambda^2\sigma_3^2(1 - \phi_3^n)]\mu \\
\beta_{1n} &= 1 - \phi_1^n \\
\beta_{2n} &= 1 - \phi_2^n \\
\gamma_n &= \frac{1-\phi_2}{\phi_2 - \phi_3}(\phi_2^{n+1} - \phi_3^{n+1}) + \frac{1}{2}\lambda^2\sigma_3^2(1 - \phi_3^n) - (1 - \phi_2^n)
\end{aligned}
$$

Appendix C: 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:

$$X_{t+1} = C + FX_t + u_{t+1} \tag{29}$$

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

$$\hat{S}_{1|0} = E(S_1). \tag{30}$$

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

$$\begin{aligned}\hat{y}_{t|t-1} &= A + BE[S_t|I_{t-1}] \\ &= A + B\hat{S}_{t|t-1}.\end{aligned} \tag{31}$$

The forcasting MSE is

$$E[(y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1})'] = BP_{t|t-1}B' + R \tag{32}$$

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

$$\begin{aligned}S_t &= \hat{S}_{t|t-1} + (S_t - \hat{S}_{t|t-1}) \\ y_t &= A + B\hat{S}_{t|t-1} + B(S_t - \hat{S}_{t|t-1}) + v_t\end{aligned} \tag{33}$$

We have the following joint distribution:

$$\begin{bmatrix} S_t \\ y_t \end{bmatrix} |I_{t-1} \sim N(\begin{bmatrix} \hat{S}_{t|t-1} \\ A + B\hat{S}_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}B' \\ BP_{t|t-1} & BP_{t|t-1}B' + R \end{bmatrix}), \tag{34}$$

Thus,

$$\begin{aligned}\hat{S}_{t|t} &\equiv E[S_t|y_t, I_{t-1}] \\ &= \hat{S}_{t|t-1} + P_{t|t-1}B'(BP_{t|t-1}B' + R)^{-1}(y_t - BS_{t|t-1} - A) \qquad (35) \\ P_{t|t} &\equiv E[(S_t - \hat{S}_{t|t})(S_t - \hat{S}_{t|t})'] \\ &= P_{t|t-1} - P_{t|t-1}B'(BP_{t|t-1}B' + R)^{-1}BP_{t|t-1} \qquad (36)\end{aligned}$$

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

$$\hat{S}_{t+1|t} = E[S_{t+1}|I_t] = F\hat{S}_{t|t} \tag{37}$$

$$\begin{aligned} P_{t+1|t} &= E[(S_{t+1} - \hat{S}_{t+1|t})(S_{t+1} - \hat{S}_{t+1|t})'] \\ &= FP_{t|t}F' + Q_t \end{aligned} \tag{38}$$

5. Maximum Likelihood Estimation of Parameters

The likelihood function can be built up recursively

$$\log L(Y_T) = \sum_{t=1}^{T} \log f(y_t|I_{t-1}), \tag{39}$$

where

$$\begin{aligned} f(y_t|I_{t-1}) &= (2\pi)^{-1/2}|H'P_{t|t-1}H + R|^{-1/2} \\ &* \ exp\{-\frac{1}{2}(y_t - A - B\hat{S}_{t|t-1})'(B'P_{t|t-1}B + R)^{-1}(y_t - A - B\hat{S}_{t|t-1})\} \\ for \quad t &= 1, 2, \ldots, T \end{aligned} \tag{40}$$

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., S. Foresi, and S. Zin (1994), Arbitrage Opportunities in Arbitrage-Free Models of Bond Pricing, unpublished, New York University.

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. 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.

6. Black, F. and P. Karasinski (1991), Bond and Option Pricing When Short Rates Are Lognormal, *Financial Analysts Journal*, (July-August), 52-59.

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 S. Schaefer, 1991, Interest Rate Volatility and the Term Structure, unpublished paper, London Business School.

9. Brown, S.J., and P.H. Dybvig, 1986, The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates, *Journal of Finance* 41:617-632.

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. Chen, L. 1996, Stochastic Mean and Stochastic Volatility – A Three-Factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management, *Financial Markets, Institutions, and Instruments*, New York University Salomon Center (March).

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

16. Duffie, D., 1992. *Dynamic Asset Pricing Theory*(Princeton University Press.)

17. Duffie, D., and R. Kan, 1993, A Yield-Factor Model of Interest Rates, Working Paper, Stanford University.

18. Eom, Y.H., 1995, Essays on Empirical Tests of Continuous Time Asset Dynamnics, unpublished, New York University.

19. 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.

20. 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.

21. 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.

22. Gong, F. and E. Remolona, 1996, Two Factors Along the Yield Curve, Federal Reserve Bank of New York.

23. Hamilton, J.D., 1994. *Time Series Analysis*, Princeton University Press.

24. 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.

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

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

27. Hull, J. and A White, 1990, Pricing Interest Rate Derivative Securities, *The Review of Financial Studies* 3, 573-592.

28. 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.

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

30. 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.

31. Lucas, Robert E., 1978, Asset Prices in an Exchange Economy, *Econometrica* 46, No.6.

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

33. Sargent, Thomas J., 1987. *Dynamic Macroeconomic Theory.* Cambridge, Ma.: Harvard University Press.

34. 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.

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

36. 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.

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

38. Vasicek, O.A., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics* 5, 177-188.

Table 1
Zero-Coupon Yields: Summary Statistics

Maturity	Monthly Observations (86:1-96:3)		
	Mean	Std Dev	1st-Auto. Corr
3 Months	5.54	1.63	0.99
6 Months	5.65	1.59	0.99
1 Year	6.26	1.69	0.98
2 Year	6.63	1.53	0.97
5 Year	7.29	1.23	0.96
10 Year	7.80	1.01	0.94

Table 2
Correlations Between Zero-Coupon Yields

Maturity	3-month	6-month	1-year	2-year	5-year	10-year
3-month	1.000					
6-month	0.995	1.000				
1-year	0.960	0.980	1.000			
2-year	0.938	0.962	0.995	1.000		
5-year	0.856	0.888	0.946	0.973	1.000	
10-year	0.739	0.774	0.849	0.894	0.969	1.000

Table 3
Parameter Estimates for the Three-factor Model Using Different Maturities

	3mon-2yr-10yr	3mon-1yr-10yr	6mon-2yr-10yr	3mon-2yr-5yr
θ	0.291 (2.482)	0.700 (2.442)	1.668 (5.165)	0.347 (7.10)
μ	10.875 (8.920)	10.521 (9.048)	10.148 (16.070)	6.900 (16.600)
ϕ_1	0.987 (0.006)	0.989 (0.007)	0.990 (0.010)	0.992 (0.008)
ϕ_2	0.849 (0.055)	0.893 (0.103)	0.876 (0.157)	0.890 (0.193)
ϕ_3	0.958 (0.052)	0.950 (0.057)	0.939 (0.140)	0.942 (0.348)
λ	-12.143 (1.496)	-12.078 (4.186)	-15.384 (2.149)	-13.122 (1.512)
σ_1	0.008 (0.044)	0.010 (0.008)	0.011 (0.003)	0.005 (0.012)
σ_2	0.116 (0.076)	0.093 (0.065)	0.132 (0.175)	0.133 (0.256)
σ_3	0.093 (0.014)	0.095 (0.026)	0.083 (0.043)	0.100 (0.087)
ρ_{12}	0.045 (2.227)	0.145 (3.857)	0.022 (0.334)	-0.020 (18.040)
ρ_{13}	-0.061 (1.165)	-0.008 (1.520)	-0.036 (1.294)	-0.120 (21.140)
ρ_{23}	0.492 (0.144)	0.727 (0.105)	0.583 (0.569)	0.715 (0.194)
e_1	0.0000 (0.0001)	0.060 (0.030)	0.004 (0.005)	0.003 (0.003)
e_2	0.377 (0.109)	0.440 (0.145)	0.443 (1.179)	0.371 (1.189)
e_3	0.509 (0.089)	0.506 (0.090)	0.522 (0.514)	0.412 (1.408)
llf	40	44	13	61

Table 4
Parameters Estimates of the Three-factor Model for Different Sample Periods Using 3-Month, 2-Year, and 10-Year Maturities

	86:1-96:3	86:1-90:12	91:1-96:3
θ	0.291 (2.482)	0.426 (34.31)	0.314 (12.16)
μ	10.875 (8.920)	8.045 (29.85)	15.041 (85.74)
ϕ_1	0.987 (0.006)	0.999 (0.004)	0.994 (0.013)
ϕ_2	0.849 (0.055)	0.895 (0.258)	0.883 (0.487)
ϕ_3	0.958 (0.052)	0.950 (0.065)	0.973 (0.212)
λ	-12.143 (1.496)	-13.970 (7.325)	-10.417 (1.879)
σ_1	0.008 (0.044)	0.009 (0.003)	0.011 (0.016)
σ_2	0.116 (0.076)	0.140 (0.207)	0.138 (0.143)
σ_3	0.093 (0.014)	0.087 (0.009)	0.124 (0.047)
ρ_{12}	0.045 (2.227)	0.018 (1.192)	0.026 (6.758)
ρ_{13}	-0.061 (1.165)	0.197 (2.276)	-0.024 (2.895)
ρ_{23}	0.492 (0.144)	0.540 (0.143)	0.763 (0.226)
e_1	0.000 (0.0001)	0.000 (0.0001)	0.000 (0.0002)
e_2	0.377 (0.109)	0.418 (0.722)	0.644 (2.514)
e_3	0.509 (0.089)	0.492 (0.325)	0.691 (2.515)
Mean Log-Likelihood	0.325	0.550	0.111

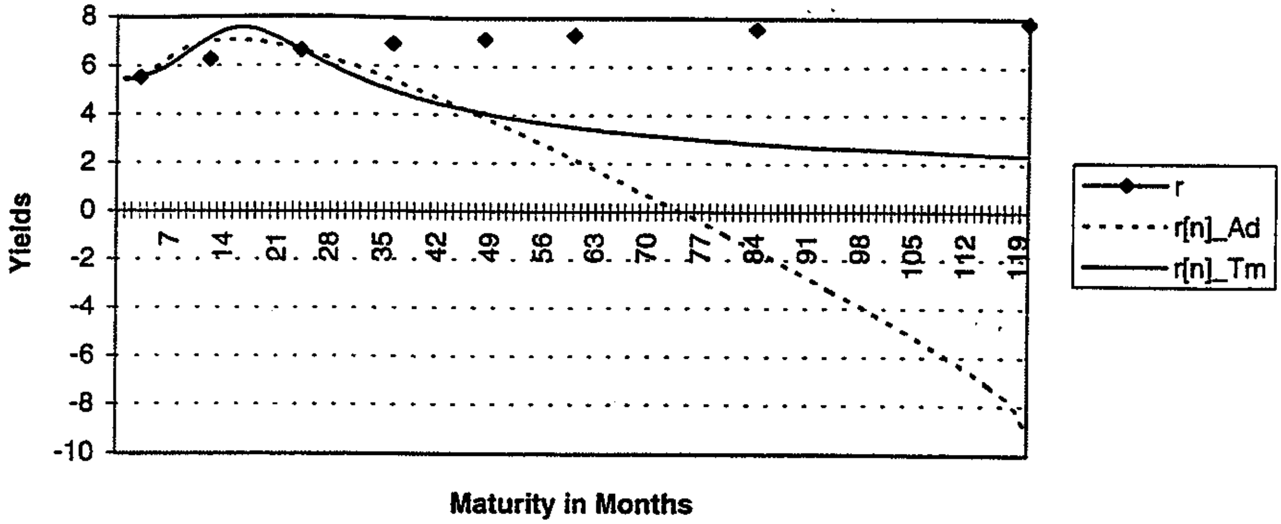

FIG. 1: Actual and Implied Yield Curves by the Two-factor Additive and Time-varying Mean Models Using 3-month and 2-year Maturities

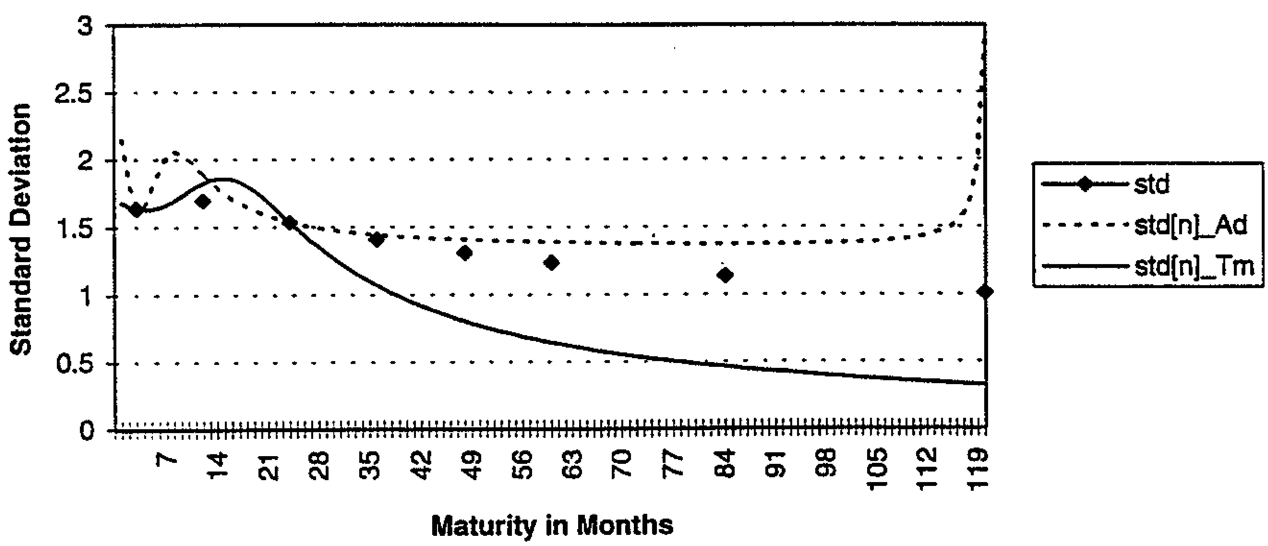

FIG. 2: Actual and Implied Volatility Curves by the Two-factor Additive and Time-varying Mean Models Using 3-month and 2-year Maturities

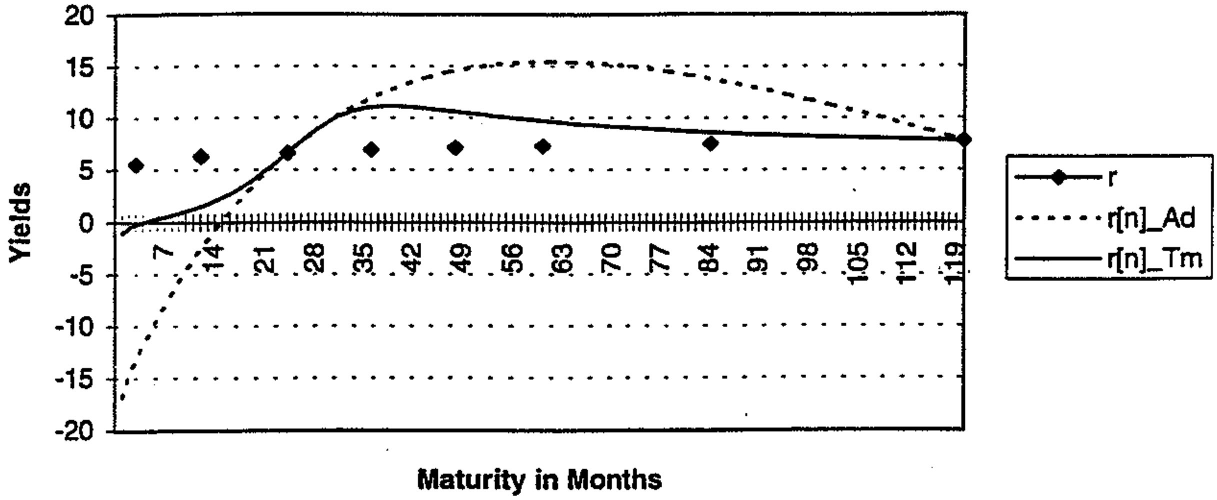

Fig.3: Implied and Actual Yield Curves by the Two-factor Additive and Time Varying Mean Models Using Maturities of 2- and 10-year Yields

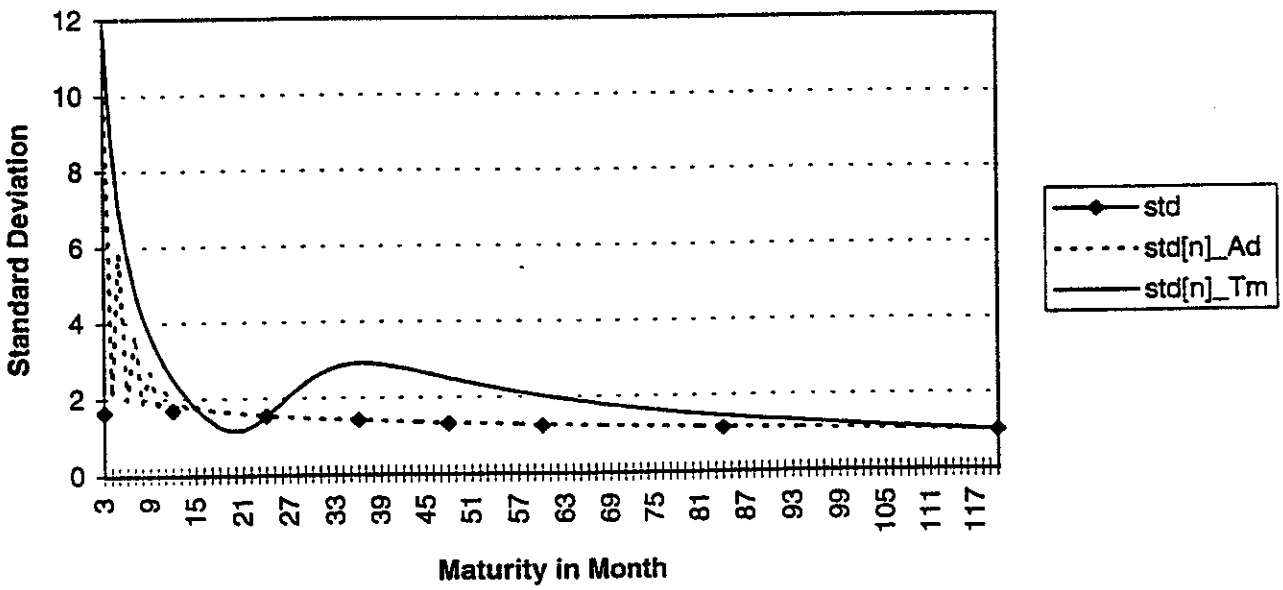

Fig. 4: Actual and Implied Volatility Curves by the Two-factor Additive and Time-varying Mean Models Using 2- and 10-year Maturities

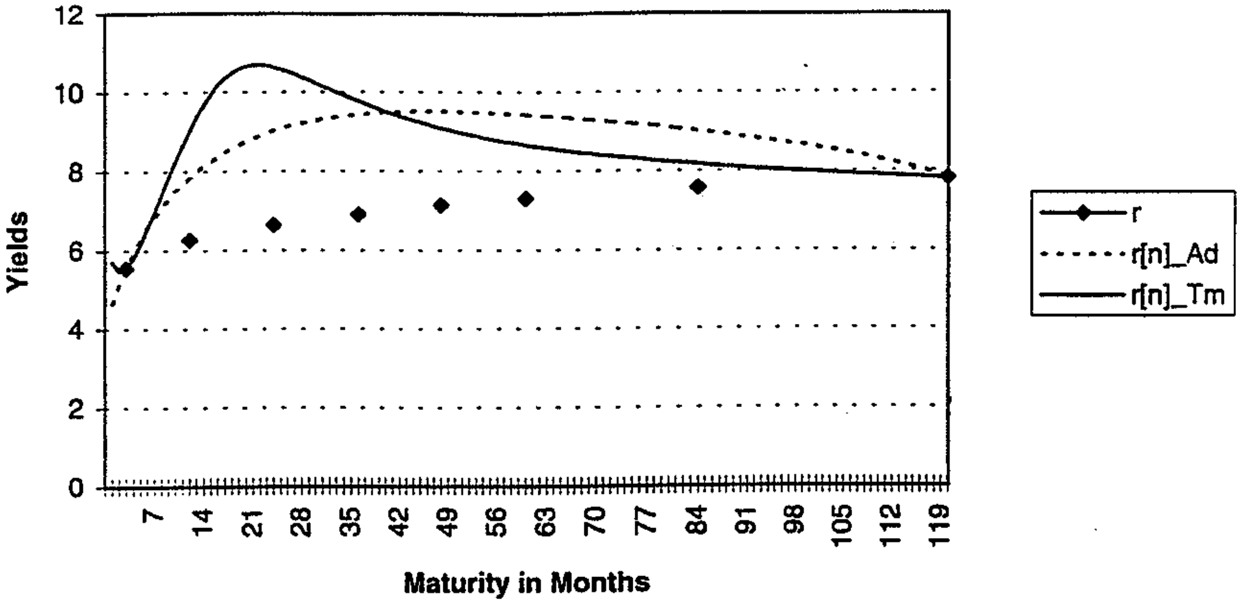

FIG.5: Actual and Implied Yield Curves by the Two-factor Additive and Time-varying Mean Models Using 3-month and 10-year Maturities

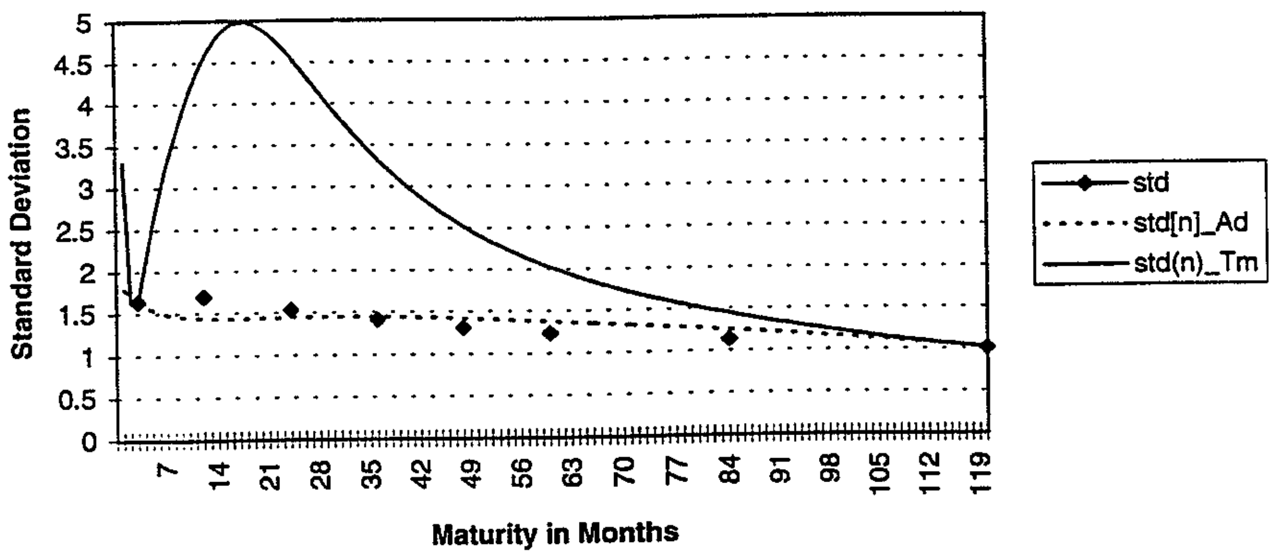

FIG. 6: Actual and Implied Volatility Curves by the Two-factor Additive and Time-varying Mean Models Using 3-month and 10-year Maturities

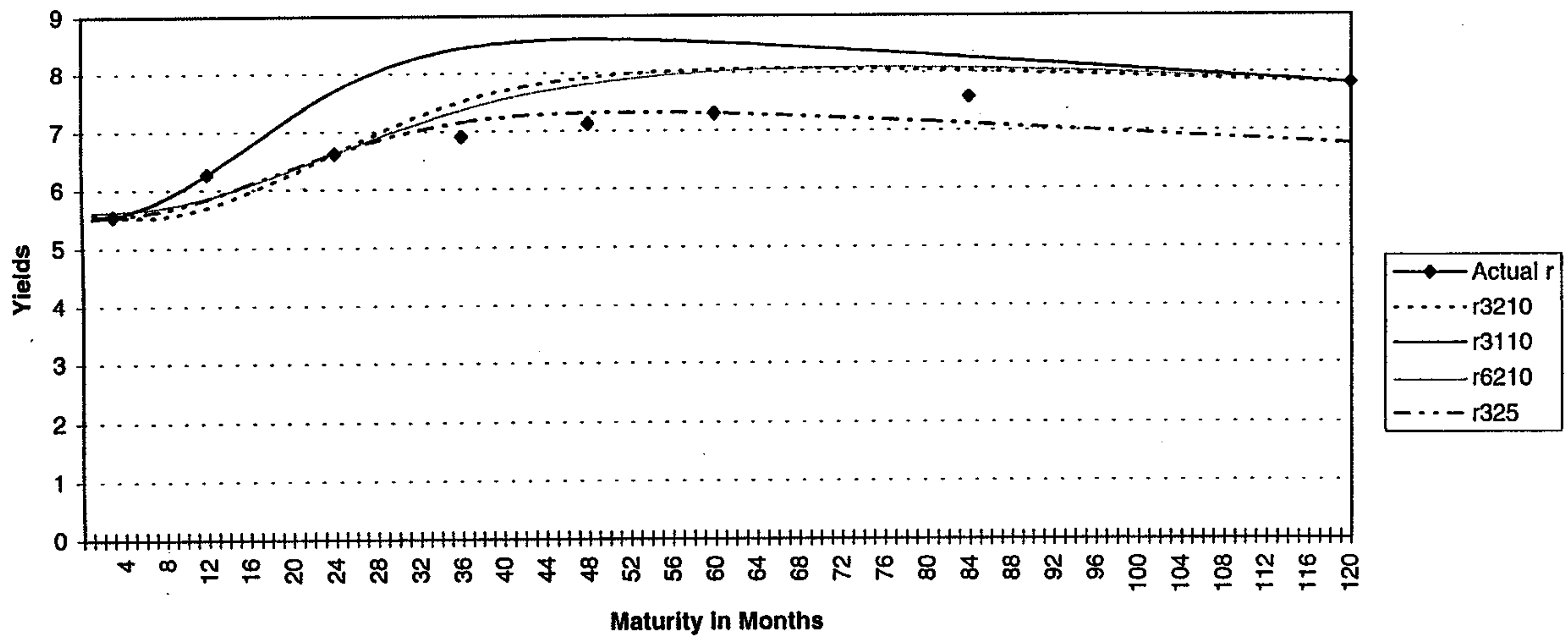

FIG. 7: Actual and Implied Yield Curves by the Three-factor Model Using Four Combinations of Three Maturities Each: 3mon, 2yr & 10yr; 3mon, 1yr & 10yr; 6mon, 2yr & 10yr; and 3mon, 2yr & 5yr

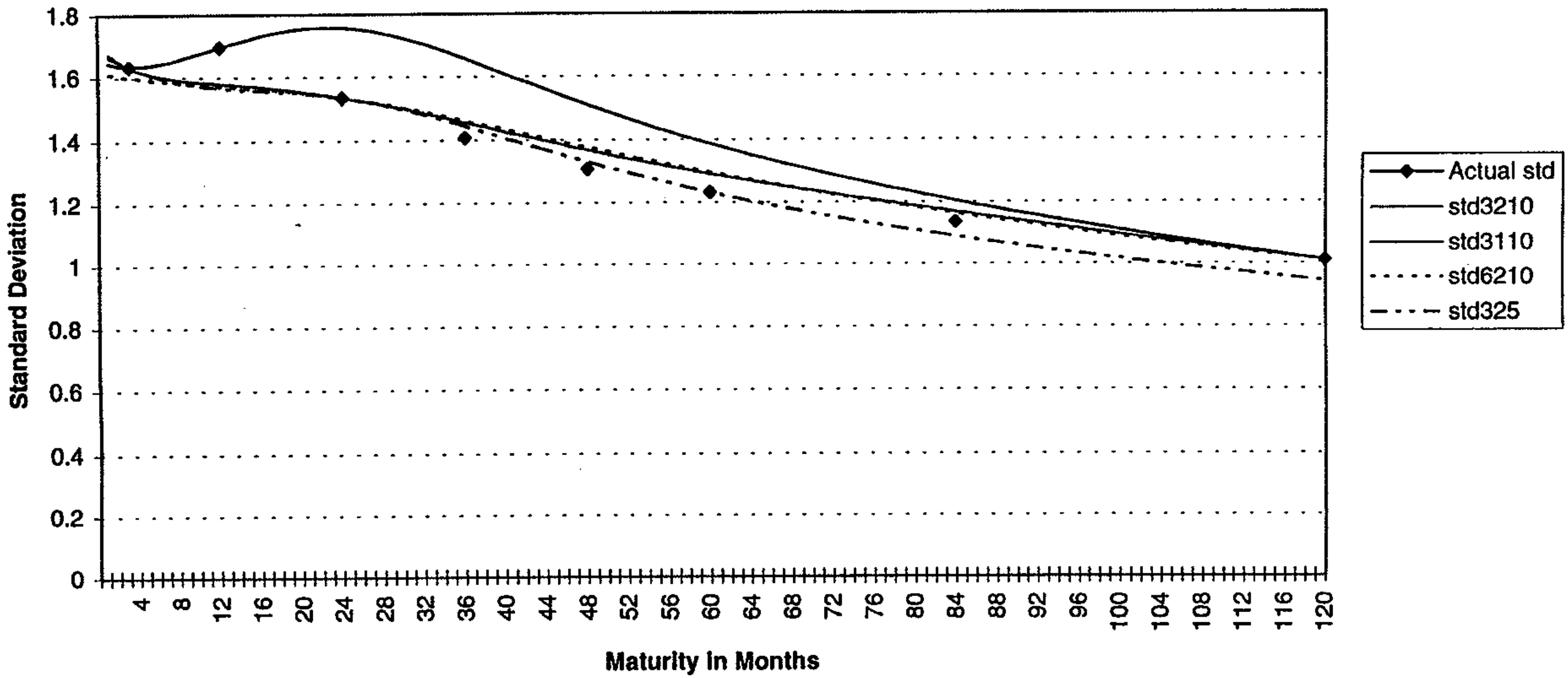

FIG. 8: Actual and Implied Volatility Curves by the Three-factor Model Using Four Combinations of Three Maturities each: 3mon, 2yr & 10yr; 3mon, 1yr & 10yr; 6mon, 2yr & 10yr; and 3mon, 2yr & 5yr

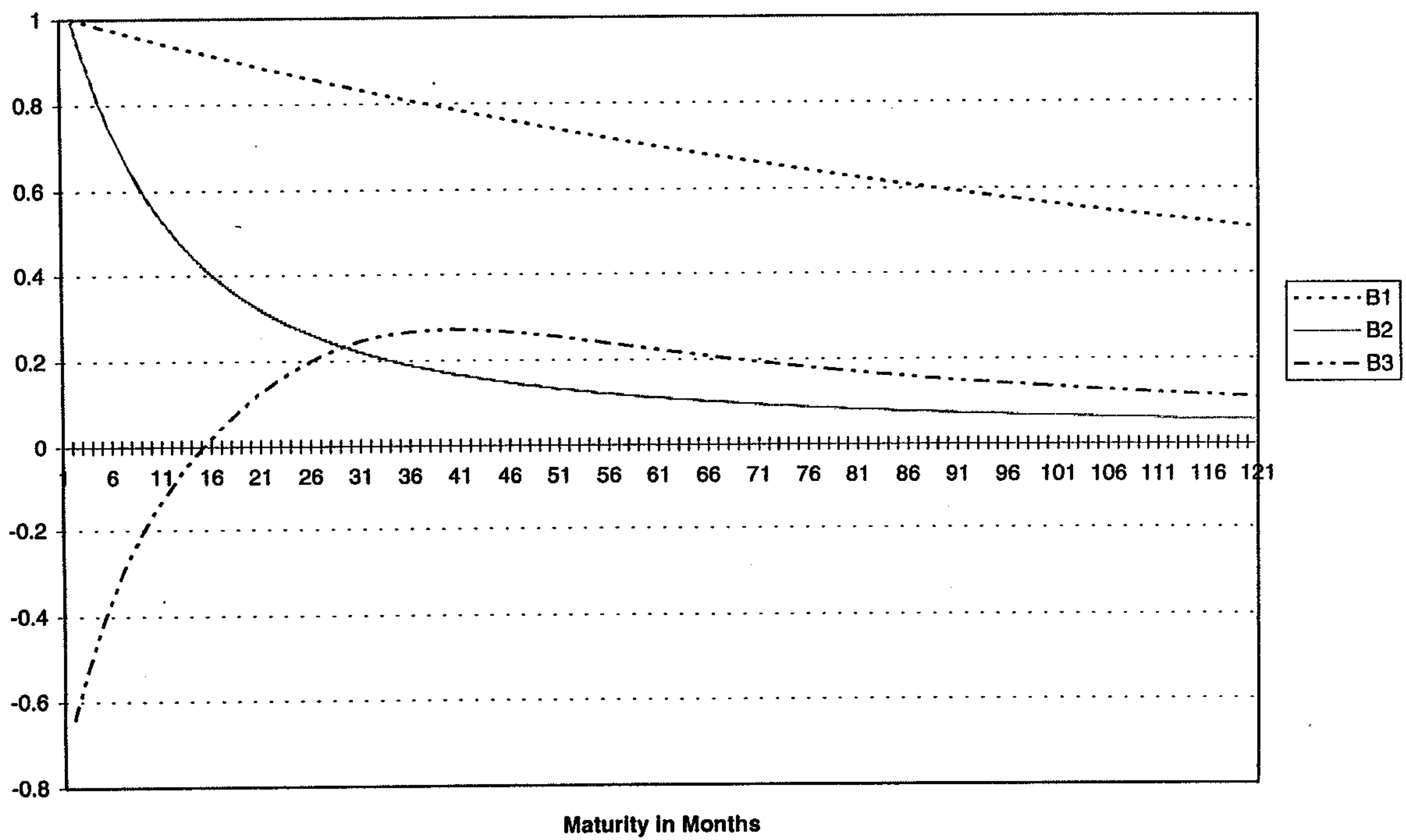

FIG. 9. Factor Loadings of Yields from the Estimated Three-Factor Model Using 3-month, 2-year and 10-year Maturities

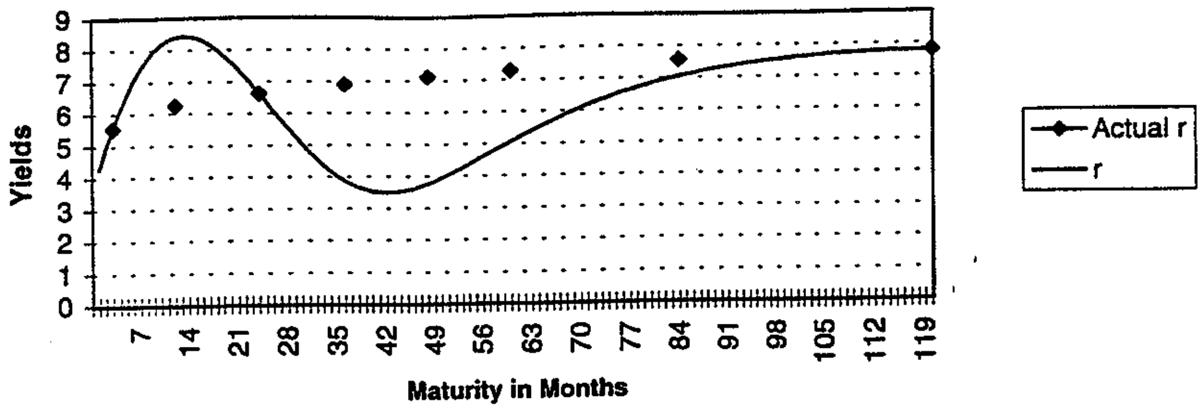

FIG. 10. Simulated Yield Curves Using Phi3=0.99 and All Other Parameters as In the Three-factor Model Estimated at 3-month, 2-year, and 10-year Maturities

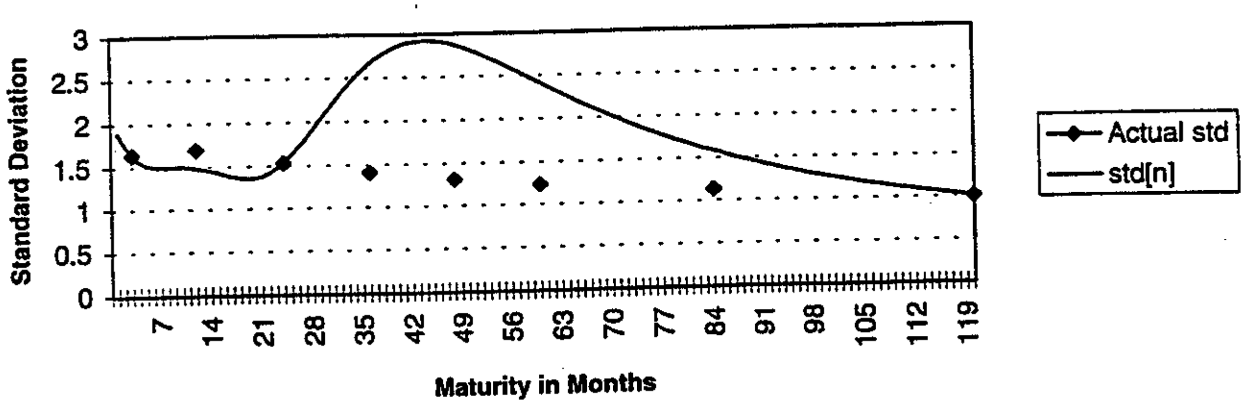

FIG. 11. Simulated Volatility Curve Using phi1=0.97, phi2=0.99 and Other Parameters in the Three-factor Model Estimated at 3-month, 2-year, and 10-year Maturities