Time Series and Cross-section Information in Affine Term-Structure Models

Frank DE JONG
Department of Finance, University of Amsterdam, Amsterdam, the Netherlands (*fdejong@fee.uva.nl*)

In this article I provide an empirical analysis of the term structure of interest rates using the affine class of term-structure models introduced by Duffie and Kan. I estimate these models by combining time series and cross-section information in a theoretically consistent way. In the estimation I use a Kalman filter based on a discretization of the continuous-time factor process and allow for a general measurement-error structure. I provide evidence that a three-factor affine model with correlated factors is able to provide an adequate fit of the cross-section and the dynamics of the term structure. The three factors can be given the usual interpretation of level, steepness, and curvature.

KEY WORDS: Kalman filter; Panel data; Term structure.

Models of the term structure of interest rates typically consist of a dynamic model for the evolution of the forcing variables, or factors, and a model for bond prices (or yields) as a function of the factors and the time to maturity. I shall refer to the former as the *time series* dimension and the latter as the *cross-section* dimension of the model. Both dimensions of the model can be analyzed separately, but there is a growing literature that estimates term-structure models using panel data—that is, combined cross-section and time series data (e.g., see Babbs and Nowman 1999; Bams and Schotman 1998; Buraschi 1996; Chen and Scott 1992; De Jong and Santa-Clara 1999; De Munnik and Schotman 1994; Duan and Simonato 1995; Frachot, Lesne, and Renault 1995; Geyer and Pichler 1997; Lund 1997; Pagan and Martin 1996; Pearson and Sun 1994). Typically, the models analyzed are multifactor versions of the Cox, Ingersoll, and Ross (CIR, 1985) model with mutually independent factors. In this article, I analyze a more general model structure, the affine class recently proposed by Duffie and Kan (1996). In this class of models the factors have an affine volatility structure, which generalizes the square-root structure of the CIR model. Moreover, the factors are allowed to be correlated. The affine term-structure model nests many well-known models, such as the one-factor Vasicek (1977) model, with constant volatility, the CIR model with square-root volatility, and the two-factor model of Longstaff and Schwartz (1992). The affine model is very tractable because interest rates are affine functions of the factors.

So far, empirical evidence on the performance of affine models is limited. Frachot et al. (1995) estimated a two-factor affine model on French data using indirect inference methods. They provided little evidence, however, on the fit of the model. Dai and Singleton (1999) estimated a general three-factor affine term-structure model on U.S. swap-rate data. They found that correlation between the factors is important for a good fit to the data. The three-factor model passes a variety of specification tests.

This article investigates a similar three-factor affine term-structure model but differs in several respects from the work of Dai and Singleton. First, my estimation methodology is based on a simple discretization of the model and does not involve computationally demanding simulation methods. In addition, I explicitly compare the empirical performance of affine models with one, two, and three factors and assess the contribution of each additional factor. A second difference is that I use a slightly broader cross-section of maturities to estimate the model, whereas Dai and Singleton used only as many maturities as factors. This is unfortunate because in that case the market price of risk parameters are not always identified. I argue that using more maturities than factors generically identifies all the parameters of the model, including the market prices of risk. In addition, using multiple maturities provides a stronger test of the restrictions imposed by the model on the cross-section of interest rates.

In the empirical analysis, I use an explicit panel-data setup. There are several advantages of using panel data. In term-structure models the absence of arbitrage opportunities imposes strong restrictions on the possible prices. Therefore, the time series model for the factors and the bond-pricing model are closely related and typically have many parameters in common. The panel-data approach fully exploits the restrictions imposed by the term-structure model and is therefore expected to give more accurate estimates of the dynamics of the term structure. Second, combined use of time series and cross-section data allows for identification of the market price of interest-rate risk, which is not identified from each dimension separately. Finally, the panel-data framework provides a natural specification test of the model by testing the restrictions imposed by the model on the parameters of the pricing equations (the cross-section dimension) and the dynamic model for the factors (the time series dimension).

A natural way to approach panel-data estimation of term-structure models is the state-space model. In the state-space model, there is a transition equation for the latent factors and a measurement equation for the interest rates on an arbitrary number of maturities. I allow for measurement error on all observed maturities and integrate out the latent fac-

Journal of Business & Economic Statistics
July 2000, Vol. 18, No. 3

tors by the Kalman filter. The prediction errors from the Kalman filter can be used to construct a quasi likelihood function, based on the conditional mean and variance of the factors. Estimation of the model is then by quasi maximum likelihood (QML), maximizing this quasi likelihood function. In this article, I show how to construct this QML estimator for the affine class of term-structure models. Because the factors are latent variables, the Kalman-filter QML estimator is not consistent, but I show by a Monte Carlo experiment that the bias is negligible for typical parameter values of the affine model.

A possible alternative to Kalman-filter QML estimation is the use of simulation-based estimators. Examples are indirect inference, applied to affine term-structure models by Frachot et al. (1995) and Buraschi (1996); the efficient method of moments, developed by Gallant and Tauchen (1996) and applied to affine term-structure models by Dai and Singleton (1999) and Pagan and Martin (1996); or Markov-chain Monte Carlo methods, used in a Bayesian analysis by Lamoureux and Witte (1998). The affine term-structure model lends itself naturally to simulation-based estimators because it is relatively easy to simulate the stochastic process for the factors. These simulation methods correct for the discretization bias and are consistent but are computationally very intensive. Given the apparent good properties of the QML estimator, simulation methods will not be used in this article.

The setup of the article is as follows. Section 1 describes the theoretical model and discusses identification issues. Section 2 discusses the empirical implementation and estimation of the model. Section 3 gives a description of the data, and Section 4 discusses the empirical results. Section 5 concludes.

1. THE AFFINE CLASS OF TERM-STRUCTURE MODELS

Endogenous term-structure models start from a process for the instantaneous short rate, r_t. Prices of zero-coupon bonds are then derived from the expected discounted payoff

$$P_t(\tau) = E_t^Q\left[\exp\left(-\int_t^T r_s\,ds\right)\right], \tag{1}$$

where the expectation is taken under the "risk-neutral" probability measure Q. Duffie and Kan (1996) proposed a class of endogenous term-structure models in which the short rate is an affine function of several underlying factors:

$$r_t = A_0 + B_0' F_t, \tag{2}$$

with $F_t \in R^n$. These factors are assumed to follow a diffusion process with an affine volatility structure

$$dF_t = \Lambda(F_t - \mu)dt + \Sigma \begin{pmatrix} \sqrt{\alpha_1 + \beta_1' F_t}\,dW_{1t} \\ \vdots \\ \sqrt{\alpha_n + \beta_n' F_t}\,dW_{nt} \end{pmatrix}, \tag{3}$$

where W_{it} are independent Wiener processes under the "real-world" or empirical probability measure P. Of course, to price bonds and other term-structure derivatives, I also need the stochastic process for the factors under the risk-neutral probability measure Q. Duffie and Kan (1996) assumed that the market price of risk for factor i is proportional to its instantaneous standard deviation, $\psi_i\sqrt{\alpha_i + \beta_i' F_t}$. Under this assumption, the transformed innovation process $dW_{it}^* \equiv dW_{it} + \psi_i\sqrt{\alpha_i + \beta_i' F_t}\,dt$ is a Wiener process under the equivalent martingale measure Q. The stochastic process for the factors under Q is given by

$$dF_t = \Lambda^*(F_t - \mu^*)dt + \Sigma \begin{pmatrix} \sqrt{\alpha_1 + \beta_1' F_t}\,dW_{1t}^* \\ \vdots \\ \sqrt{\alpha_n + \beta_n' F_t}\,dW_{nt}^* \end{pmatrix}. \tag{4}$$

The risk-neutral intercept and mean-reversion parameters are related to the parameters of the real-world dynamics through

$$\Lambda^* = \Lambda - \Sigma\Psi\mathcal{B}' \tag{5}$$

and

$$\Lambda^*\mu^* = \Lambda\mu + \Sigma\Psi\alpha, \tag{6}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)', \mathcal{B} = (\beta_1, \ldots, \beta_n)$, and $\Psi = \operatorname{diag}(\psi_1, \ldots, \psi_n)$.

Duffie and Kan (1996) showed that in the affine model the price of a zero-coupon bond with time to maturity τ is an exponential affine function of the vector of factors,

$$P_t(\tau) = \exp[-A(\tau) - B(\tau)' F_t]. \tag{7}$$

Due to this form, the interest rates or yields on zero-coupon bonds are a linear function of the factors, where the intercept and factor loadings are time-invariant functions of the time to maturity

$$Y_t(\tau) \equiv -\ln P_t(\tau)/\tau = A(\tau)/\tau + B(\tau)'/\tau \cdot F_t. \tag{8}$$

The coefficients $A(\tau)$ and $B(\tau)$ satisfy the system of ordinary differential equations

$$\frac{dA(\tau)}{d\tau} = A_0 - (\Lambda^*\mu^*)' B(\tau) - \frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)a_{ij} \tag{9}$$

and

$$\frac{dB(\tau)}{d\tau} = B_0 + (\Lambda^*)' B(\tau) - \frac{1}{2}\sum_i\sum_j B_i(\tau)B_j(\tau)b_{ij}, \tag{10}$$

where the scalars a_{ij} and the vectors b_{ij} are defined by $a_{ij} + b_{ij}' x \equiv [\Sigma \operatorname{diag}(\alpha + \mathcal{B}' x)\Sigma']_{ij}$.

The affine model contains several well-known models as special cases. The model of Langetieg (1980), which generalizes the Vasicek (1977) model to more dimensions, is obtained if $\mathcal{B} = 0$. The generalized Cox et al. (1985) model is obtained if $\alpha = 0$ and $\mathcal{B}$ is diagonal. In the latter model, all yields are guaranteed to be positive (see Pang and Hodges 1996). If, in addition, the mean-reversion matrix Λ and the correlation matrix Σ are diagonal, the factors follow mutually independent stochastic processes and I obtain a two-factor CIR model that is observationally equivalent to the Longstaff and Schwartz (1992) model. Jegadeesh and Pennacchi (1996) proposed a model in which the short rate fluctuates around a stochastic mean. This model is also a special case of the affine class with a particular recursive structure for Λ.

Empirically, not all the parameters of the affine model can be identified and certain normalizations are necessary. The first identification issue concerns the "intercepts" of the model. As long as A_0 is unrestricted, the mean of the factors, μ, and the intercept of the variances, α, are not separately identified. We therefore normalize $\mu = 0$. With this normalization, A_0 is the mean of the instantaneous short rate under the P measure, and the vector α can be interpreted as the average volatility of the factors.

Pang and Hodges (1996) showed that bond prices are invariant under invertible transformations of the factors. The first implication of this result is that bond prices are invariant under scale transformations of the factors. Hence, without loss of generality, I normalize $B_0 = \iota$ so that the instantaneous interest rate equals a constant plus the sum of the factors ($r_t = A_0 + \iota' F_t$). The second implication of this result is that only the product matrix $\mathcal{K} \equiv \Sigma^{-1}\Lambda\Sigma$ can be identified. Without loss of generality, I could therefore assume, as Dai and Singleton (1999) did, that Σ is the identity matrix. A more convenient normalization for the empirical work is that Λ is diagonal, and the diagonal elements of Σ are equal to 1. I parameterize $\Lambda = \operatorname{diag}(-\kappa_1, \ldots, -\kappa_n)$. For any parameterization, the vector $\kappa = (\kappa_1, \ldots, \kappa_n)'$ consists of (minus) the eigenvalues of the mean-reversion matrix. These eigenvalues are always identified and independent of the normalization. Finally, the existence conditions of Duffie and Kan (1996) for the affine model, discussed by Dai and Singleton (1999), impose some additional restrictions on Σ and $\mathcal{B}$. In particular, when treating the off-diagonal elements of Σ as free parameters, $\mathcal{B}'\Sigma$ should be a diagonal matrix. This condition fixes all the off-diagonal elements of $\mathcal{B}$. We parameterize $\mathcal{B}'\Sigma = \tilde{\beta}$, where $\tilde{\beta}$ is diagonal and treat $(\tilde{\beta}_{11}, \ldots, \tilde{\beta}_{nn})$ as free parameters. This parameterization of the model can be summarized in the following system, that is equivalent to Equations (2) and (3):

$$r_t = A_0 + \iota' F_t \tag{11}$$

and

$$dF_t = \Lambda F_t dt + \Sigma \begin{pmatrix} \sqrt{\alpha_1 + \tilde{\beta}_{11}(\Sigma^{-1}F_t)_1}\, dW_{1t} \\ \vdots \\ \sqrt{\alpha_n + \tilde{\beta}_{nn}(\Sigma^{-1}F_t)_n}\, dW_{nt} \end{pmatrix}, \tag{12}$$

with $\Lambda = \operatorname{diag}(-\kappa_1, \ldots, -\kappa_n)$. Except for the diagonal elements of Σ that are fixed at 1, all parameters in this affine system can be identified. The identification of the market price of risk parameters, ψ, will be discussed separately in Section 2.

An additional identification problem arises in the special case of the multivariate Vasicek model with constant variances ($\mathcal{B} = 0$). Dai and Singleton (1999) showed that the Vasicek model is invariant under so-called unitary rotations of the factors. Essentially this means that only $n(n+1)/2$ elements in $\mathcal{K}$ can be identified. To identify the model, I assume, in addition to the previous normalizations, that Σ is upper triangular.

2. EMPIRICAL IMPLEMENTATION OF THE AFFINE MODEL

The simplest approach to estimating an n factor model is to select n yields with different maturities and to obtain the factors by "inverting" the model. In some special cases, in particular the Gaussian multifactor Vasicek or Langetieg (1980) models and in CIR models with uncorrelated factors, the discrete transition density of the factors is known. Multiplying this density with the Jacobian of the transformation gives the exact likelihood function. This is the approach followed by Chen and Scott (1993) and Pearson and Sun (1994) in an analysis of two-factor CIR models. The choice of maturities to construct the factors is rather arbitrary, however, and the results of the model will depend on the choice. Another limitation of the exact ML approach is that it is not easily generalized to models with correlated factors because, except in the Gaussian case, the multivariate transition density of the factors is unknown. More seriously, only using as many maturities as factors neglects potentially useful information in the other maturities. Instead, our estimation is based on the state-space form of the model.

Let there be observations for maturities τ_1 through τ_k. Collect the observed zero-coupon bond yields for period t and the coefficients in the vectors y_t and A and matrix B, defined as

$$y_t \equiv \begin{pmatrix} Y_t(\tau_1) \\ \vdots \\ Y_t(\tau_k) \end{pmatrix}, \qquad A \equiv \begin{pmatrix} A(\tau_1)/\tau_1 \\ \vdots \\ A(\tau_k)/\tau_k \end{pmatrix},$$

$$B \equiv \begin{pmatrix} B(\tau_1)'/\tau_1 \\ \vdots \\ B(\tau_k)'/\tau_k \end{pmatrix}.$$

The state-space form of the model is

$$y_t = A + BF_t \tag{13}$$

and

$$F_{t+h} = \Phi F_t + \nu_{t+h}, \tag{14}$$

where h is the time interval between two observations. The first equation in this system is the measurement equation. The coefficients of the measurement equation are functions of the parameters under the risk-neutral distribution, $(A_0, \mu^*, \Lambda^*, \alpha, \tilde{\beta}, \Sigma)$. Due to the model structure, the parameters of the risk-neutral distribution depend on the basic parameters $(A_0, \kappa, \alpha, \tilde{\beta}, \Sigma, \psi)$ via the restrictions (5) and (6). The second equation in the state-space model is the transition equation; it is the discrete-time equivalent of Equation (3) with the normalization $\mu = 0$ imposed. The parameters of the transition equation follow from the conditional mean and variance of the factors: $E[F_{t+h}|F_t] = \Phi F_t$ and $\operatorname{var}(\nu_{t+h}) = \operatorname{var}(F_{t+h}|F_t) \equiv q(F_t)$. Appendix A shows that the conditional mean and variance are affine functions of the current level of the factors and depend on the parameters $(\kappa, \alpha, \tilde{\beta}, \Sigma)$.

The state-space model gives a clear insight on the identification of the market prices of risk. The mean of the factors under the risk-neutral distribution, μ^*, is identified from the time series average of the vector of observed yields, provided that the dimension of the vector of yields, y_t, is at least $n+1$, with n being the number of factors. The additional yield is necessary to identify the intercept of the equation for the short rate, A_0. Because I fix the mean under P of the factors by imposing $\mu = 0$, the parameters ψ are exactly identified from Relation (6). In contrast, in the models of Pearson and Sun (1994) and Dai and Singleton (1999), who estimated n-factor affine models using n series of zero-coupon yields, identification of all market prices of risk is achieved via Restriction (5), but this only leads to identification if the volatility depends on the level of the factors. In the constant volatility model, one element of μ^* and hence ψ is not identified if only n zero-coupon bond yields are used for estimation.

The state-space form also highlights that the panel structure imposes overidentifying restrictions on the model, in particular the equality of the parameters $(\kappa, \alpha, \tilde{\beta}, \Sigma)$ in the cross-section and time series dimension. These restrictions provide a natural specification test of the affine model, similar to cross-equation tests used in models of rational expectations.

The affine model predicts the exact relation $y_t = A + BF_t$ between the factors and the interest rates. When using more maturities than factors, this exact relation cannot be satisfied by all elements of the yield vector. Therefore, some form of measurement error is necessary, but the theoretical model is silent about the covariance structure of the errors. Several assumptions have been made in the literature. Chen and Scott (1993), for a model with two factors and four maturities, assumed that two yields are observed without error so that the model for these two maturities can be inverted to obtain the factors. The other yields, or linear combinations thereof, are assumed to be measured with a normally distributed measurement error. Several works—for example, Duan and Simonato (1995), Geyer and Pichler (1995), and Jegadeesh and Pennacchi (1996)—assumed that all interest rates are observed with some measurement error, which is both serially and cross-sectionally uncorrelated. De Jong and Santa-Clara (1999) had a similar econometric approach for estimating a Markovian model of the Heath, Jarrow, and Morton (1992) type. Bams and Schotman (1998) followed a similar approach but allowed for some correlation between the errors for different maturities. Lund (1997) and Frachot et al. (1995) pointed out that a diagonal error covariance matrix is not robust under linear transformations of the data. They proposed using a more general, nondiagonal, cross-sectional correlation matrix for the measurement errors.

Following these suggestions, I assume that the measurement errors have zero mean and are serially uncorrelated but may be cross-sectionally correlated with time-invariant covariance matrix H. The most convenient way to parameterize H is as LDL', where L is lower triangular with ones on the diagonal and D is the diagonal matrix of eigenvalues. This form makes H positive definite by construction. This parameterization is more general than the assumptions made in many other works, and it makes the estimates of the parameters invariant to linear transformations of the yield vector y_t. The full state-space form of the model with measurement errors is then

$$y_t = A + BF_t + e_t, \qquad \mathrm{var}(e_t) = H, \tag{15}$$

and

$$F_{t+h} = \Phi F_t + \nu_{t+h}, \qquad \mathrm{var}(\nu_{t+h}) = q(F_t). \tag{16}$$

Given this state-space setup, the most convenient way to estimate the parameters is by QML based on the Kalman filter. The relevant equations for the Kalman filter in the affine term-structure model are given in Appendix B. The Kalman-filter QML estimator is consistent and efficient if the factors and the error terms follow normal distributions. In most affine term-structure models, the conditional distribution of the factors is not normal, but if the conditional mean and variance of the factors are correctly specified, one could expect the estimates obtained from the Kalman filter to be consistent by the QML principle.

There is one subtle problem with this argument that arises because the conditional variance of the factors depends on the current value of the factors, which are latent variables that can be estimated but not observed exactly. Therefore, the conditional variance used in the likelihood function will not be correct. An additional problem is that the conditional distribution of the factors is not normal, which invalidates the updating rules in the Kalman filter. As a result, the QML estimates obtained from the Kalman filter will be inconsistent [for an extensive discussion of this point see Lund (1997)]. The inconsistency could be removed by applying the indirect inference methods of Gouriéroux, Montfort, and Renault (1992) or the efficient method of moments of Gallant and Tauchen (1996). This is the approach followed by Frachot et al. (1995), Pagan and Martin (1996), and Dai and Singleton (1999). Lamoureux and Witte (1998) proposed a Bayesian approach in which the posterior distribution of the parameters is obtained by draws from the exact conditional distribution of the factors and the measurement errors. Their procedure, however, is numerically very intensive.

To assess the bias in the Kalman-filter QML estimator, I performed a small Monte Carlo experiment. The way the data are generated is as follows. First, the latent factor process is simulated from the diffusion (3) using an Euler discretization scheme with 25 intermediate steps per month. Each month the value of the factor was recorded and yields for maturities 3 months and 1, 5, and 10 years were calculated from the model equation (8). Then I generate measurement errors from the normal distribution for each observation. To keep the model simple, the measurement-error covariance matrix is assumed to be of the form $H = h^2 I$ so that the errors for different maturities have the same variance and are cross-sectionally and serially uncorrelated. Each simulation generates a sample of monthly observations on four different maturities. Finally, in each simulation the model parameters $(A_0, \kappa, \alpha, \beta, h)$ are estimated using the Kalman-filter QML estimator. In the estimation, the market price of the risk parameter is treated as known. The reason for this is that I want to concentrate on the effects of the discretization of the transition equation, in which the market price of risk does not play a role.

I run this Monte Carlo experiment for two sets of parameters. The first set is based on the estimates of the one-factor models in Table 3, Section 3. In the second experiment, I pick the parameters from the estimates of the second factor in the two-factor model. This factor has a stronger mean reversion, a higher variance and a stronger level effect in the volatility function. The hypothetical true parameter values and several descriptive statistics of the Monte Carlo estimates are reported in Table 1, for 100 replications and for two sample sizes, 300 and 1,200 observations. The main conclusion I can draw from this table is that the QML estimator has no systematic bias. The only significant overestimation occurs in the second experiment for the mean-reversion parameter, but the bias is fairly small. This evidence confirms the results of Lund (1997), who suggested that, for parameters typically found in estimates of term-structure models, the bias in the QML estimator is not particularly large. Therefore, I refrain from using computationally intensive simulation-based estimation techniques and report the QML estimates.

3. DATA

My database is the extended McCulloch dataset (McCulloch and Kwon 1993). This dataset contains monthly observations on U.S. interest rates with maturities running from 1 month to 30 years. The original data series starts in 1947 and ends in 1991. The data are zero-coupon rates that were calculated from prices of coupon bonds using McCulloch's interpolation method. From 1985 only bonds that do not have prepayment provisions are used; earlier data may include such bonds. The interpolation has some relatively large standard errors for data from the fifties and sixties. Moreover, there are a lot of missing observations in the early part of the sample. For these reasons, I work with a subsample of the data that starts in January 1970 and ends in February 1991. In total, there are 254 monthly observations.

As for the choice of maturities, I picked the maturities for inclusion in the yield vector as follows. For the maturities over 10 years, the bond data are quite scarce, so the interpolation is not very accurate. Another problem shows up in the very short-term interest rates. The one- and two-month interest-rate series show some exceptionally large one-period changes. I feel more confident using interest rates with maturities of 3 months and longer. Because I have a full error covariance matrix, the number of parameters to be estimated is potentially large. To keep the estimation feasible, I confine myself to four maturities—3 months, 1 year, 5 years, and 10 years.

Figure 1 graphs the data, and Table 2 gives some descriptive statistics. The long-maturity interest rates are somewhat

Table 1. Monte Carlo Results

Parameter	True value	Median	Mean	St. dev.	t value
		300 observations			
A_0	6.4642	6.4612	6.4337	.4617	−.6606
κ	.0601	.0601	.0599	.0100	−.2018
α	1.0468	1.0528	1.0527	.0786	.7495
β	.3961	.3906	.3909	.0843	−.6141
h	.2479	.2475	.2478	.0047	−.1546
		1,200 observations			
A_0	6.4642	6.4755	6.4545	.1484	−.6516
κ	.9601	.0606	.0605	.0037	1.0599
α	1.0468	1.0455	1.0458	.0216	−.4670
β	.3961	.3962	.3979	.0284	.6273
h	.2479	.2472	.2474	.0026	−1.6976
		300 observations			
A_0	6.1000	6.0903	6.0923	.0897	−.8597
κ	1.3056	1.3225	1.3135	.1083	.7305
α	4.5400	4.5334	4.5825	.4333	.9800
β	2.2960	2.3710	2.3311	.4895	.7169
h	.2479	.2479	.2481	.0059	.4496
		1,200 observations			
A_0	6.1000	6.0891	6.0569	.3337	−1.2920
κ	1.3056	1.3260	1.3296	.0624	3.8489
α	4.5400	4.6095	4.7600	1.9051	1.1548
β	2.2960	2.3796	2.3840	.5062	1.7382
h	.2479	.2481	.2976	.4144	1.2006

NOTE: This table reports results of a Monte Carlo experiment for the QML estimator of a one-factor affine model, with $r_t = A_0 + F_t$, $dF(t) = -\kappa F_t dt + \sqrt{\alpha + \beta F_t} dW(t)$. The number of simulation runs is 100. In each simulation, 300 or 1,200 monthly observations on 3-month, 1-, 5-, and 10-year interest rates were generated according to the model. Measurement error with diagonal covariance matrix with standard deviation h was added to the observations.

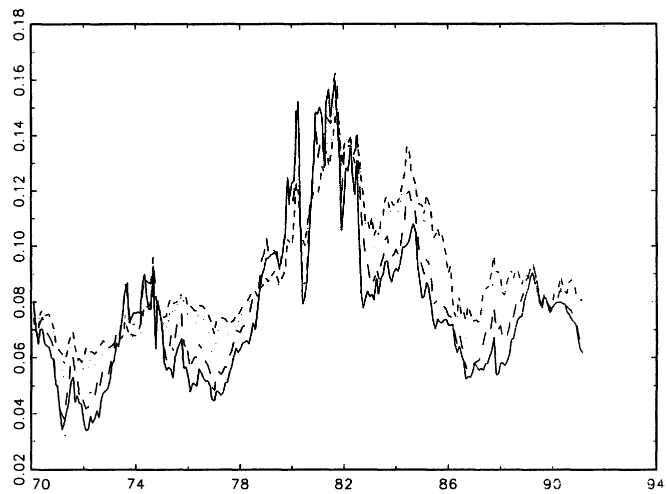

Figure 1. U.S. Term-Structure Data. The figure shows the 3-month, 1-, 5-, and 10-year zero coupon yields constructed by McCulloch and Kwon (1993) from U.S. treasury coupon bonds: ———, 3 months; – – –, 1 year; · · · ·, 5 years; —·· —, 10 years.

Table 2. Descriptive Statistics of the McCulloch and Kwon Data

	Maturity			
	3-month	1-year	5-year	10-year
Mean	7.68	8.20	8.75	8.95
Standard deviation	2.68	2.58	2.27	2.13
Minimum	3.38	3.74	5.15	5.72
Maximum	16.00	16.35	15.70	15.07

NOTE: The data consist of 254 monthly U.S. interest rates, from January 1970 to February 1991. The interest rates are expressed in percent per annum. Source of the data is McCulloch and Kwon (1993).

Table 3. Estimation Results of the One-Factor Affine Models

Model	A_0 (×100)	κ	$\tilde{\alpha}$ (×10^4)	β (×100)	ψ (×10^{-2})	κ^*	2 ln L
Affine	6.4642	.0601	−1.5137	.3961	−.1481	.0014	710.45
	(.2465)	(.0074)	(.4369)	(.0637)	(.0069)	[495.10]	
CIR	5.8099	.0429		.2168	−.1446	.0116	702.43
	(.5397)	(.0042)		(.0231)	(.0110)	[60.00]	
Vasicek	7.3146	.0222	1.9980		−.0928	.0222	677.60
	(1.4600)	(.0028)	(.2284)		(.0178)	[31.19]	

NOTE: This table reports QML estimates and standard errors for the parameters of one-factor term-structure models with short rate process $dr(t) = \kappa(\theta - r(t))dt + \sqrt{\tilde{\alpha} + \beta r(t)}dW(t)$. The table also reports the mean-reversion parameter κ^* and the half-life of the factors, [ln(2)/κ^*], under the risk-neutral distribution.

less variable than the short rates. Moreover, on average the term structure is upward sloping. The large volatilities of the interest rates around 1980 show up clearly. Because this is also a period with high levels of interest rates, the data give some intuitive support for models in which the conditional variance depends on the level of the interest rates.

4. EMPIRICAL RESULTS

One of the main issues in term-structure modeling is the number of factors to include. In this section, I therefore discuss the empirical results on one-, two-, and three-factor affine models. For each model I provide an extensive specification analysis. Estimation of all models is by the Kalman-filter QML estimator described in Section 2. Heteroscedasticity-consistent standard errors are calculated by the method of White (1982).

4.1 One-Factor Models

In this section I discuss the results of modeling the term structure by one-factor affine term-structure models. The one-factor version of the affine model is very tractable because the differential equations (9)–(10) have analytical solutions for $A(\tau)$ and $B(\tau)$. In the one-factor model, Equations (3) and (4) specialize to

$$dr_t = \kappa(\theta - r_t)dt + \sqrt{\tilde{\alpha} + \beta r_t}dW_t \tag{17}$$

and

$$dr_t = \kappa^*(\theta^* - r_t)dt + \sqrt{\tilde{\alpha} + \beta r_t}dW_t^*, \tag{18}$$

where I follow the usual convention in one-factor models to specify the process for the instantaneous short rate. Of course, this is equivalent with my general model by defining the factor as $F_t = r_t - \theta$ and $\tilde{\alpha} = \alpha - \beta\theta$. The term structure is given by $P_t(\tau) = \exp[-\tilde{A}(\tau) - B(\tau)r_t]$, where the functions $\tilde{A}(\tau)$ and $B(\tau)$ can be found from a straightforward generalization of the standard CIR equations that were given, for example, by Hull (1993):

$$\tilde{A}(\tau) = -\frac{2\tilde{\phi}}{\beta}\ln\left(\frac{2\gamma e^{[(\kappa^*+\gamma)/2]\tau}}{(\kappa^*+\gamma)(e^{\gamma}-1)+2\gamma}\right) + \frac{\tilde{\alpha}}{\beta}(\tau - B(\tau))$$

$$B(\tau) = \frac{2(e^{\gamma\tau}-1)}{(\kappa^*+\gamma)(e^{\gamma\tau}-1)+2\gamma},$$

$$\kappa^* = \kappa + \psi\beta, \quad \tilde{\phi} \equiv \kappa\left(\theta + \frac{\tilde{\alpha}}{\beta}\right), \quad \gamma \equiv \sqrt{(\kappa^*)^2 + 2\beta}.$$

For the special case of the Vasicek model ($\beta = 0$), the coefficients are

$$\tilde{A}(\tau) = R_\infty(\tau - B(\tau)) + \frac{\tilde{\alpha}}{4\kappa}B(\tau)^2$$

$$B(\tau) = \frac{1-\exp(-\kappa\tau)}{\kappa},$$

where $R_\infty \equiv \theta - (\psi\tilde{\alpha}/\kappa) - (\tilde{\alpha}/2\kappa^2)$ is the yield on infinite maturity bonds.

The parameters to be estimated are the mean-reversion coefficient κ, the long-run mean of the short rate θ, the variance parameters $\tilde{\alpha}$ and β, and the market price of risk parameter ψ. In Table 3 we report estimates of the one-factor affine model and two special cases, the CIR model ($\tilde{\alpha} = 0$) and the Vasicek model ($\beta = 0$).

The estimated mean-reversion coefficient is very small, around .06. The implied mean reversion under the risk-neutral distribution is even slower: κ^* in the affine model is .0014, which implies a half life, $\ln(2)/\kappa^*$, of around 500 years. The result in the CIR model is virtually the same, and in the Vasicek model the estimated half-life is around 30 years. This slow mean reversion implies very flat-fitted term structures. Although the infinite maturity yield must be constant if κ^* is positive, the mean reversion is slow enough to create considerable and almost parallel movements in, say, 10-year rates.

The estimated intercept of the instantaneous variance $\tilde{\alpha}$ is negative in the affine model. The average volatility, $\alpha = \tilde{\alpha} + \beta\theta$, is positive, however, but lower than the estimated constant volatility parameter for the Vasicek model. The "slope" coefficient β is positive and significant, and the constant volatility assumption ($\beta = 0$) of the Vasicek model is clearly rejected. Interestingly, the estimate of β in the affine model is larger than the comparable estimate for the CIR model, where $\tilde{\alpha}$ is restricted to be 0. The sensitivity of the conditional variance to the level of the short rate is therefore stronger in the affine model than in the CIR model. Time-series-based studies have reported a similar phenomenon (e.g., see Chan, Karalyi, Longstaff, and Sanders 1992). The estimates of the market price of risk parameters are significantly negative and of the same order of magnitude in all specifications. This result implies that the risk premium for holding long-term bonds is positive (the risk premia are derived in Appendix C). The implied risk premium for a 10-year bond is around 1.65%

annually, which corresponds quite well to the observed risk premium.

Given the estimated parameters, I can construct residuals of the model, defined as the difference between the observed yields and the predicted yields and therefore equal to the prediction errors generated by the Kalman filter. In a well-specified model, the time series average of residuals should be close to 0 for all maturities. In addition, the residuals should be serially uncorrelated. Table 4 provides some summary statistics on the residuals of the one-factor affine model. The model on average overestimates short-maturity interest rates and slightly underestimates interest rates for longer maturities. The standard deviations of the residuals are also very large, especially for short maturities: The standard deviation is around 150 basis points for the three-month rate. In addition, there is strong residual serial correlation. The first-order autocorrelations are around .9 or higher, the twelfth-order autocorrelations are around .5.

The bad fit of the one-factor model is illustrated in Figure 2, which graphs the average of the fitted and observed term structures for all maturities available in the McCulloch and Kwon (1993) dataset, from 1 month to 10 years. The fitted term structure was calculated from the smoothed estimates of the factors, $F_{t|T}$ [the smoothing equations of Hamilton (1994, chap. 13.6) were used to calculate these], and then calculating the fitted yields as $y_{t|T}(\tau) = A(\tau) + B(\tau)F_{t|T}$ for all values of τ in the specified range. The graph also plots the root mean squared error of the difference between the observed and fitted yields. The figure shows that the model fits the long end pretty well but fails to capture the short end of the yield curve.

Another way to judge the quality of the model is by regressing the observed yields of all maturities between 1 month and 10 years on a constant and the estimated (smoothed) factors. Because the data are close to being nonstationary, this regression is done in first differences. The regression coefficients should be more or less the same as the factor loadings for these maturities obtained from the term-structure model. Figure 3 shows that this holds for the long maturities (over 5 years), but for shorter maturities there are large differences between the estimated sensitivities and the model values.

Table 4. Residuals of the One-Factor Affine Model

	Maturity			
	.25	1	5	10
Mean	.7027	−.2395	.0760	.1146
Standard deviation	1.5967	1.3009	.5513	.3851
ρ_1	.8957	.8459	.4427	.2031
ρ_{12}	.4635	.4265	.2070	.1443
Correlation matrix				
	1.0000	.9500	.6958	.4123
	.9500	1.0000	.8230	.5525
	.6958	.8230	1.0000	.8731
	.4123	.5525	.8731	1.0000

NOTE: This table reports descriptive statistics (sample mean, standard deviation, and serial correlations, ρ_k) of the residuals of the one-factor affine term-structure model. The scale of the residuals is percentage points.

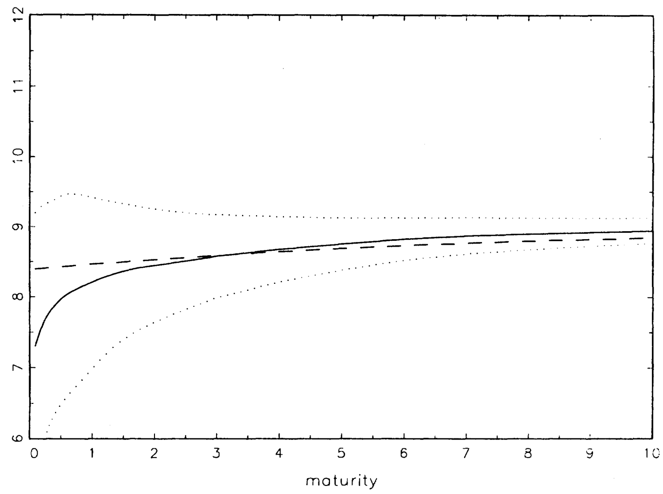

Figure 2. Fit of the One-Factor Affine Model. The figure shows the average actual and fitted term structures, as well as the root mean squared error of the difference between fitted and observed values, in the one-factor affine model: ———, average yield curve; – –, fitted curve; · · · ·, RMSE; · · · ·, bands.

A more formal way to test the specification of the model is by testing the restrictions the model imposes between the parameters of the pricing equations (the cross-section dimension) and the parameters of the time series dimension. In addition to the parameters of the standard model, the following additional parameters for the time series dimension can be identified—a mean-reversion parameter $\hat{\kappa}$ and variance parameters $\hat{\alpha}$ and $\hat{\beta}$. Table 5 reports results for the one-factor model with separate coefficients for the cross-section and the time series dimension. The restrictions between the time series and cross-section parameters are rejected by the likelihood ratio (LR) test. The LR test statistic is 17.76, which is larger than the 5% critical value of a chi-squared distribution with 3 df. The key to this rejection is that the time series estimates show a much stronger mean reversion and higher instantaneous variance than the cross-section estimates. To fit the rather flat shape of the

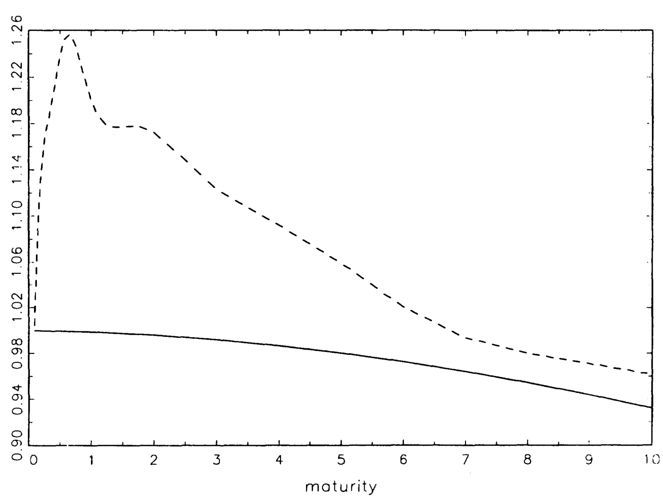

Figure 3. Regression of Observed Yields on Fitted Factors. The figure shows the coefficients of a regression (in first differences) of the observed time series of yields on the time series of fitted factors in the one-factor affine model.

Table 5. Estimation Results for One-Factor Affine Models With Separate Time Series Parameters

Model	A_0 (×100)	κ	$\tilde{\alpha}$ (×10^4)	β (×100)	ψ (×10^{-2})	κ^*	$\hat{\kappa}$	$\hat{\alpha}$ (×10^4)	$\hat{\beta}$ (×100)	2 ln L
Affine	7.4027	.0230	6.0774	.0088	−.0630	.0225	.2183	−1.7248	.4318	728.21
	(.8220)	(.0034)	(1.1225)	(.0002)	(.0021)	[31.39]	(.1116)	(.4813)	(.0788)	
CIR	7.5788	.0307		.5943	−.0688	−.0102	.2041		.1988	714.55
	(1.0065)	(.0049)		(.0907)	(.0068)		(.1242)		(.0232)	
Vasicek	7.5574	.0218	6.0548		−.0631	.0218	.2018	1.9220		693.54
	(1.0509)	(.0031)	(.8864)		(.0040)	[31.70]	(.1321)	(.2185)		

NOTE: This table reports QML estimates and standard errors of one-factor affine term-structure models with separate parameters for the cross-section and time series dimension. Parameters with a ^ are the time series parameters. See also the note to Table 3.

observed yield curves, a slow mean reversion is necessary, whereas in the time series dimension the mean reversion of interest rates is quite strong. The estimated mean reversion under the risk-neutral distribution is comparable to the previous estimates for the affine and Vasicek models. For the CIR model, the estimate of κ^* is negative, but for the maturities I consider, the shape of the $B(\tau)$ curve is not very different from the other models.

All these results point at substantial misspecification of the one-factor affine term-structure model. In particular, the model fails to give a good fit of the term structure at the short end. Moreover, the dynamics of the yield curve are not well described as evidenced by the strong residual serial correlation and the differences in parameter estimates for the time series and cross-section dimensions. I therefore now turn to an analysis of multifactor models.

4.2 Two-Factor Models

In this section I present estimates of affine term-structure models with two factors. From the discussion on identification in Section 2, it follows that the most general two-factor model has, as free parameters, (minus) the diagonal elements of Λ, denoted by κ; the off-diagonal elements of Σ; the intercept A_0; the variance parameters α and $\tilde{\beta}$; and the market prices of risk ψ. In addition to the most general model, I estimate some restricted versions of the affine model. The first special case is an affine model with uncorrelated factors, which imposes that Σ is the identity matrix. This model is equivalent to the two-factor CIR model with an intercept proposed by Pearson and Sun (1994). The other special case I consider is the generalized Vasicek model with constant volatilities ($\hat{\beta} = 0$). I also present estimates of the Vasicek model with uncorrelated factors.

The estimation results in Table 6 show that there are two factors with very different properties. The mean reversion of the first factor is similar to the mean reversion in the one-factor model, with a half-life over 30 years. The second factor shows a much stronger mean reversion with a half-life of less than one year. The solid lines in Figure 4 graph the factor loadings $B(\tau)$. The first factor loading, $B_1(\tau)$, is very flat; the impact of a shock in the first factor is around 1 for all maturities considered. Theoretically, $B_1(\tau)$ should converge to 0 for large τ, but apparently this convergence is so slow that it is hardly detectable at horizons up to 10 years. So, although the model implies constant infinite maturity yields, long-run yields can vary substantially. The factor loading of the second factor declines much faster but is not negligible even for the longest maturities I consider.

To interpret the factors, we graph the fitted factors together with functions of the data in Figure 5. The first panel graphs the first factor (with the estimated intercept A_0 added) and the 10-year interest rate. The two series are very highly correlated, and therefore this figure suggests an interpretation of the first factor as the level of the yield curve. The gap between the level of the factor and the

Table 6. Estimation Results of the Two-Factor Affine Models

Model	A_0 (×100)	κ_i	α_i (×10^4)	β_{ii} (×100)	σ_{ij}	ψ_i (×10^{-2})	κ_i^*	2 ln L
Affine	6.10	.0341	.51	.3043	.0470	−.0655	.0054	1,376.40
	(.88)	(.0074)	(.28)	(.0548)	(.0222)	(.0172)	[127.65]	
		1.3056	4.54	2.2960	−.2688	−.2061	.8411	
		(.1126)	(.71)	(.4908)	(.0737)	(.0095)	[.82]	
Affine	10.75	.0164	2.48	.4261		−.0454	−.0031	1,333.93
(Σ = 0)	(.17)	(.0045)	(1.00)	(.1357)		(.0016)		
		1.1103	5.86	1.8889		−.1449	.8369	
		(.0765)	(.68)	(.3674)		(.0025)	[.83]	
Vasicek	11.77	.0234	1.58		−.0118	−.0213	.0234	1,225.31
	(1.34)	(.0029)	(.20)		(.0260)	(.0128)	[29.59]	
		.8424	6.46			−.1378	.8424	
		(.0510)	(1.28)			(.0167)	[.82]	
Vasicek	10.80	.0232	1.52			−.0275	.0232	1,225.11
(Σ = 0)	(.82)	(.0029)	(.19)			(.0102)	[29.85]	
		.8469	6.51			−.1558	.8469	
		(.0498)	(1.17)			(.0153)	[.82]	

NOTE: This table reports QML estimates and standard errors of two-factor affine term-structure models with $r_t = A_0 + \iota' F_t$, $dF_t = \Lambda F_t dt + \Sigma S(t) dW_t$, $\Lambda_{ii} \equiv -\kappa_i$, $S(t)_{ii} \equiv \alpha_i + \tilde{\beta}_{ii}(\Sigma^{-1}F_t)_i$, and ψ the vector of market prices of risk. Diagonal elements of Σ are fixed at 1. The table also reports the eigenvalues, κ_i^*, of the mean reversion matrix under the risk-neutral distribution, Λ^*, and the associated half-lives.

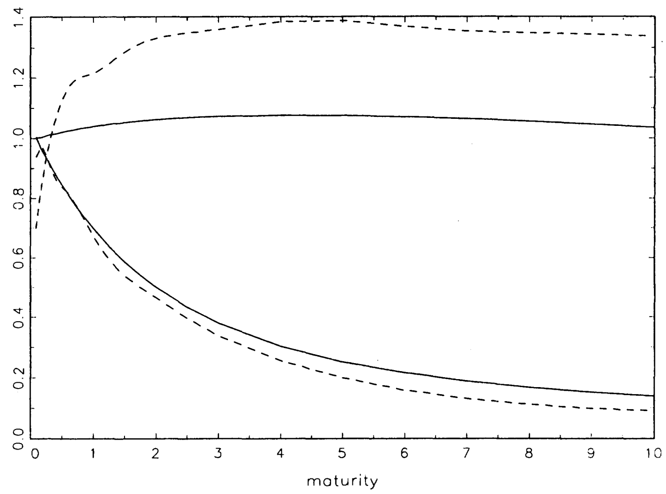

Figure 4. Regression of Observed Yields on Fitted Factors. The figure shows the implied factor loadings of the two-factor model (solid lines) the coefficients of a regression (in first differences) of the observed time series of yields on the time series of fitted factors in the two-factor affine model (dashed lines).

Table 7. Residuals of the Two-Factor Affine Model

	Maturity			
	.25	*1*	*5*	*10*
Mean	.0106	.2331	.1462	.1157
Standard deviation	.7462	.7603	.5086	.4042
ρ_1	.1851	.4629	.4178	.3524
ρ_{12}	−.1087	.1782	.1982	.1876
Correlation matrix				
	1.0000	.8345	.6828	.5990
	.8345	1.0000	.8910	.7900
	.6828	.8910	1.0000	.9267
	.5990	.7900	.9267	1.0000

NOTE: This table reports descriptive statistics (sample mean, standard deviation, and serial correlations, ρ_k) of the residuals of the two-factor affine term-structure model. The scale of the residuals is percentage points.

10-year rate is approximately equal to the risk premium on 10-year bonds of around 1.65%. The second factor is very closely related to the spread between the three-month rate and the 10-year rate. These results are in line with the results of Litterman and Scheinkman (1991), who documented a level factor and a slope factor in a principal-components decomposition of the term structure. Balduzzi, Das, and Foresi (1995) modeled the second factor as a central tendency, which is quite similar to a time-varying slope.

A strong result of my analysis is the significant correlation between the factors, which is evident from the nonzero off-diagonal elements of Σ. Indeed, the LR test strongly rejects independence of the factors: The LR test statistic for the restrictions $\sigma_{12} = 0$ and $\sigma_{21} = 0$ is 42.47, larger than the 5% critical value of the $X^2(2)$ distribution. To see how this correlation affects the dynamics of the factors, I write the model in feedback form, with mean reversion matrix $\mathcal{K} = \Sigma^{-1}\Lambda\Sigma$. With the numbers from Table 6, the mean-

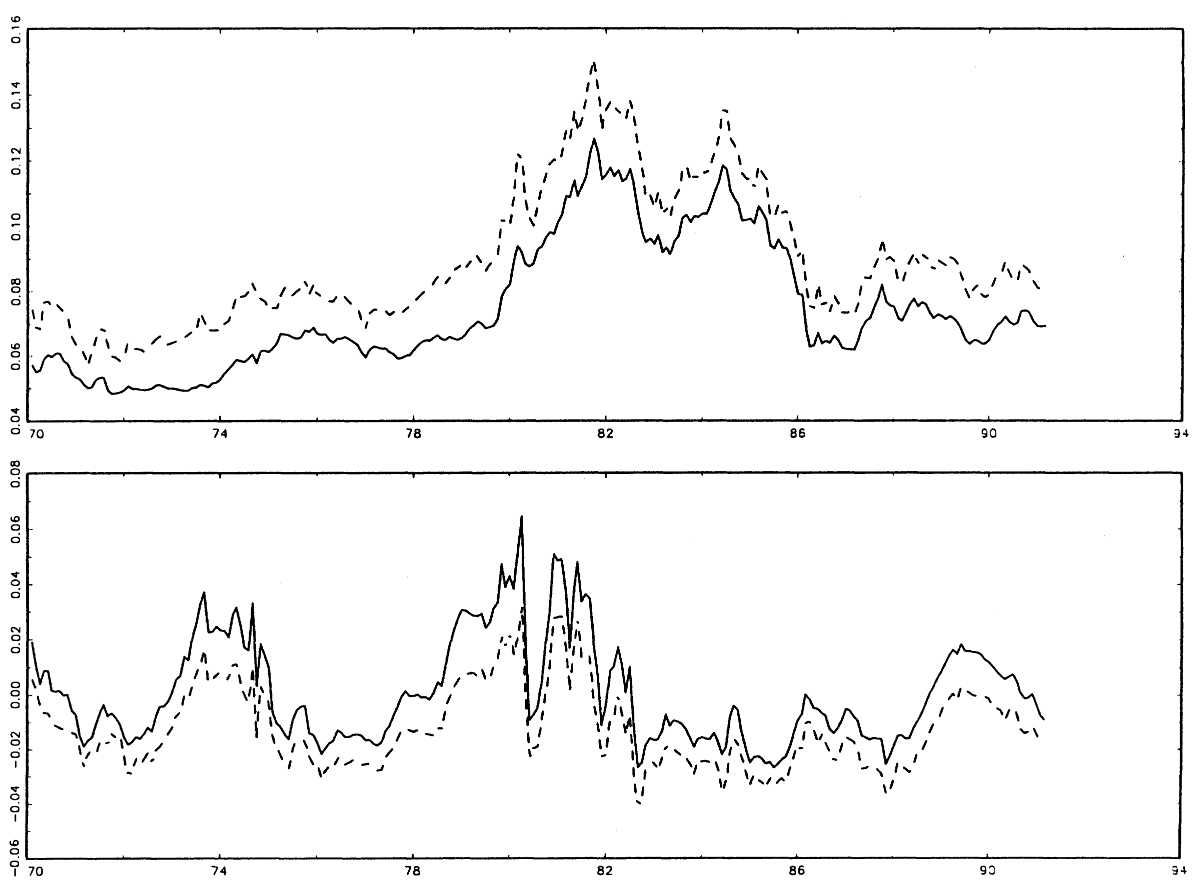

Figure 5. Fitted Factors in the Two-Factor Affine Model. The figure shows the estimated factors in the two-factor affine model.

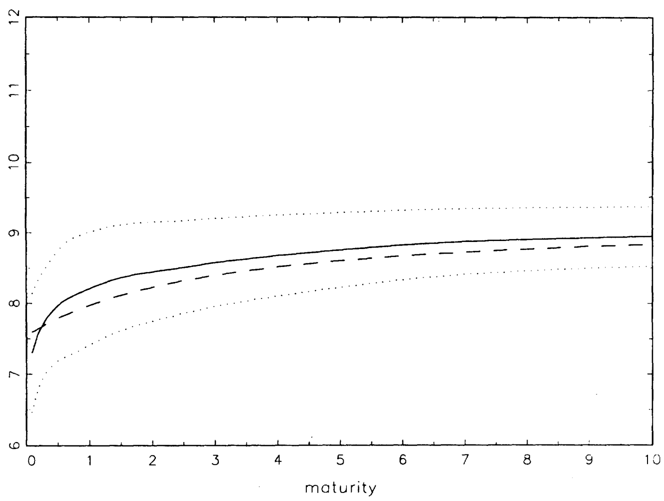

Figure 6. Fit of the Two-Factor Affine Model. The figure shows the average actual and fitted term structures, as well as the root mean squared error of the difference between fitted and observed values, in the two-factor affine model: ———, average yield; – –, fitted curve; · · · ·, RMSE; · · · ·, bands.

reversion matrix in feedback form is

$$\mathcal{K} = \begin{pmatrix} -.0499 & .0590 \\ .3375 & -1.2898 \end{pmatrix}.$$

This specification highlights that there is two-way feedback between the factors. For example, a positive innovation in the second factor has a positive impact on the first factor, and vice versa. Notice that the magnitude of the parameters should be related to the volatility of the factors; the conditional variance of the second factor is about 10 times the conditional variance of the first factor.

Turning to the variance parameters, there is a significant level effect in the volatilities of both factors. The estimates of $\tilde{\beta}_{11}$ and $\tilde{\beta}_{22}$ are strongly significant, and the restriction that they are 0 (the Vasicek model) is strongly rejected. The market prices of risk for both factors are negative, which implies a positive risk premium for holding long-term bonds. With the instantaneous variance evaluated at the long-term mean of the factors, the implied risk premium for a 10-year bond is 1.64%, split over the first factor (.34%) and the second factor (1.30%).

The two-factor affine model provides a substantially better fit of the term-structure data than the one-factor model. The variance of the residuals is smaller than in the one-factor model and, more importantly, the serial correlation is substantially lower (see Table 7). Figure 6 graphs the average fitted term structure. An inspection of the graph reveals that most of the improved fit is for the three-month rate. The model still overestimates interest rates for maturities of one year and longer. So, although the two-factor model improves on the one-factor model, the steepness of the observed term structures at the short end is not yet fully captured by the two-factor model.

Figure 4 shows the fitted slope coefficients of a time series regression of the observed yields on the estimated factors. The estimated coefficients of the first factor, with a slow mean reversion, are close to the factor loadings implied by the model. For the second factor, however, there are substantial differences between the estimated sensitivities and the factor loadings implied by the model. Again, this is probably due to the bad fit of the model in the very short end of the term structure.

Finally, I formally tested the specification of the two-factor affine model by allowing for a different set of parameters for the cross-section and the time series dimension of the model. To avoid overparameterization, I keep Σ restricted and only estimate separate time series parameters for the mean reversion ($\hat{\kappa}$) and volatility ($\hat{\alpha}$ and $\hat{\beta}$). Table 8 reports the results for this less restricted model. The parameter restrictions implied by the theoretical model are rejected on a 5% significance level; the LR test statistic for the equality of $(\kappa, \alpha, \hat{\beta})$ in the cross-section and time series dimension is 92.88, much larger than the 5% critical value of a $X^2(6)$ distribution. Just like in the one-factor model, the main difference between the two sets of estimates is the strength of the mean reversion of the factors. The half-life of the first factor in the time series dimension is around 6 years, compared with around 30 years in the cross-section dimension. The conditional variances of the factors are also very different. As a result, the model restrictions on the time series parameters are rejected.

4.3 A Three-Factor Affine Model

Although the two-factor model provides a substantial improvement over the one-factor model, there is still some misspecification, especially in the short end, where the observed yield curves are on average steeper than the curves generated by the model. Andersen and Lund (1997) and Dai and Singleton (1999) argued that a third factor is necessary to fully describe the term structure. Although the general

Table 8. Estimation Results for Two-Factor Affine Models With Separate Time Series Parameters

Model	A_0 (×100)	κ_i	α_i (×10^4)	$\tilde{\beta}_{ii}$ (×100)	σ_{ij}	ψ_i (×10^{-2})	κ_i^*	$\hat{\kappa}_i$	$\hat{\alpha}_i$ (×10^4)	$\hat{\beta}_{ii}$ (×100)	2 ln L
Affine	16.08	.0008	.02	.3067	.1450	−.0151	−.0834	.0195	3.51	.2969	1,469.28
	(8.72)	(.0002)	(.01)	(.1642)	(.0377)	(.0111)		(.0047)	(1.50)	(.1028)	
		1.6441	263.81	69.9727	−.1804	−.0162	.5904	.9222	6.53	1.8824	
		(1.0468)	(160.94)	(50.7893)	(.1622)	(.0372)	[1.17]	(.2296)	(6.59)	(.3629)	
Vasicek	9.79	.0246	.00		−.0288	−.0272	.0246	.0953	1.79		1,299.73
	(.39)	(.0030)	(.00)		(.0284)	(.0086)	[28.19]	(.0680)	(.22)		
		.7520	262.61			−.0110	.7520	.8944	6.31		
		(.0392)	(30.23)			(.0020)	[.92]	(.4896)	(1.40)		

NOTE: This table reports QML estimates and standard errors of two-factor affine term-structure models with separate time series coefficients $\hat{\kappa}_i$, $\hat{\alpha}_i$, and $\hat{\beta}_i$. The model estimated is $r_t = A_0 + \kappa' F_t$, $dF_t = \Lambda F_t dt + \Sigma S(t) dW_t$, $\Lambda_{ii} \equiv -\kappa_i$, $S(t)_{ii} \equiv \alpha_i + \tilde{\beta}_{ii}(\Sigma^{-1} F_t)_i$, and ψ is the vector of market prices of risk. Diagonal elements of Σ are fixed at 1. The table also reports the eigenvalues, κ_i^*, of the mean-reversion matrix under the risk-neutral distribution, Λ^*, and the associated half-lives.

Table 9. Estimation Results Three-Factor Affine Model

A_0	κ_i	α_i	$\tilde{\beta}_{ii}$	σ_{1i}	σ_{2i}	σ_{3i}	ψ_i	κ_i^*	$2 \ln L$
11.27	.0549	2.48	.3498	1.0000	.0325	−.0460	.1214	−.0012	1,602.96
(14.37)	(.0412)	(6.30)	(.3151)		(.0257)	(.1376)	(.0427)		
	1.7407	18.66	3.2055	.3433	1.0000	−.6806	−.2179	.8981	
	(1.1920)	(26.87)	(2.6778)	(1.7716)		(.3892)	(.0668)	[.77]	
	3.2117	34.98	.0013	−.8137	−.3818	1.0000	−.2786	3.3693	
	(.9569)	(31.51)	(.0129)	(2.7473)	(.6113)		(.0469)	[.21]	

NOTE: This table reports QML estimates and standard errors of the three-factor affine term-structure model $r_t = A_0 + \kappa' F_t$, $dF_t = \Lambda F_t dt + \Sigma S(t) dW_t$, $\Lambda_{ii} \equiv -\kappa_i$, $S(t)_{ii} \equiv \alpha_i + \tilde{\beta}_{ii}\Sigma^{-1}F_{ti}$, and ψ is the vector of market prices of risk. Diagonal elements of Σ are fixed at 1. The table also reports the eigenvalues, κ_i^*, of the mean-reversion matrix under the risk-neutral distribution, Λ^*, and the associated half-lives.

affine three-factor model has many parameters (e.g., Σ has six free parameters), convergence of the estimator was reasonably quick.

The estimation results are presented in Table 9. The first two factors are similar to the factors of the two-factor affine model. The third factor is very different, however. The mean reversion of this factor is very strong, the implied half-life is around three months, and the instantaneous volatility is much higher than the volatility of the other factors. The correlation with the two other factors is negative. This is most easily judged from the feedback matrix of the three-factor model:

$$\mathcal{K} = \begin{pmatrix} .0429 & .0754 & -.1293 \\ 1.6079 & -1.2028 & -1.3894 \\ 3.2622 & .8283 & -3.8474 \end{pmatrix}.$$

This matrix shows that the impact of a shock in the third factor on the other factors is negative. On the other hand, the impact of shocks in the first two factors on the third factor is positive.

The graph of the fitted factors in Figure 7 confirms the interpretation of the first and second factor as level and slope factors, although the second factor is somewhat smoother than the slope of the term structure. For an interpretation of the third factor, I plot a measure of the curvature of the term structure in the graph of the third factor. The curvature is defined as two times the one-year interest rate, minus the sum of the three-month rate and the 10-year rate. The third factor in the affine model is closely related to this curvature measure, although it is somewhat smoother. This result is consistent with the results of Dai and Singleton (1999), who found that their third factor is closely related to the curvature of the term structure. They also documented a

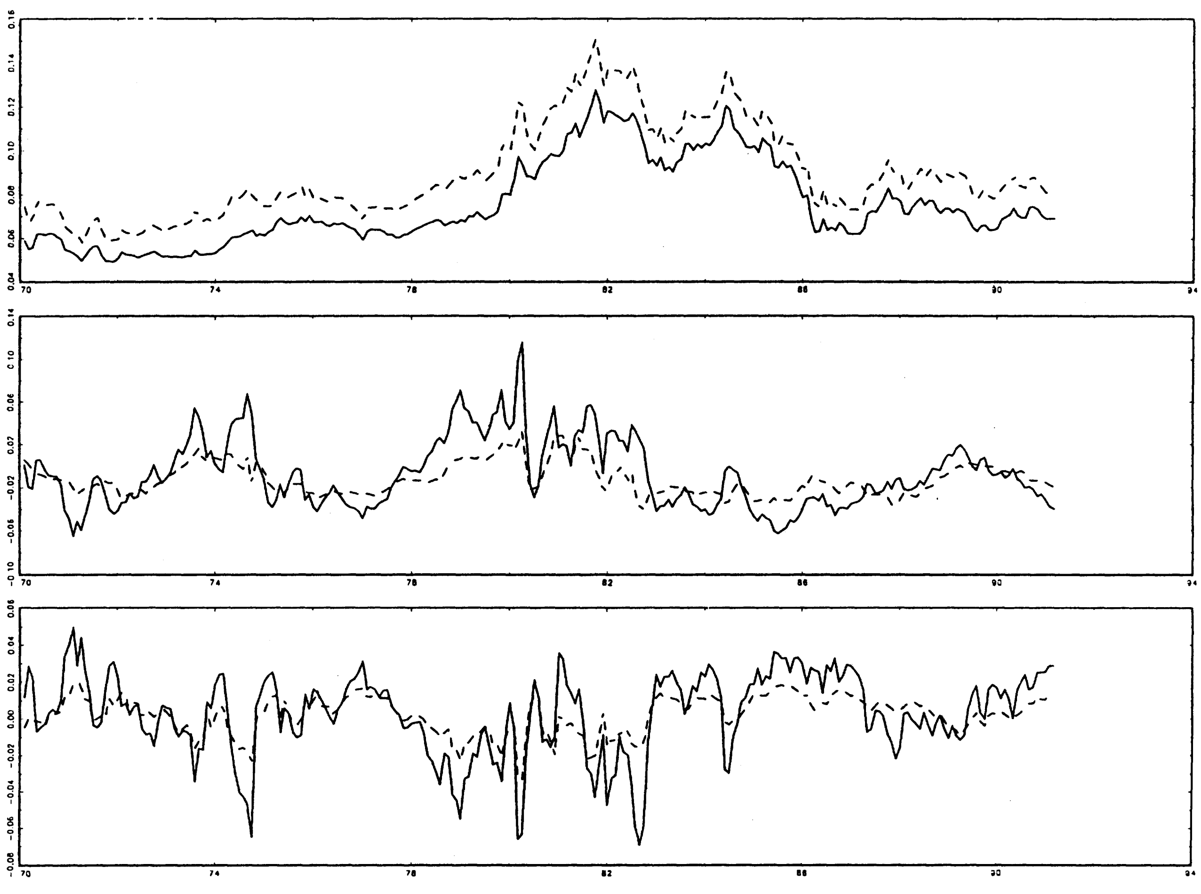

Figure 7. Fitted Factors in the Three-Factor Affine Model. The figure shows the estimated factors in the three-factor affine model.

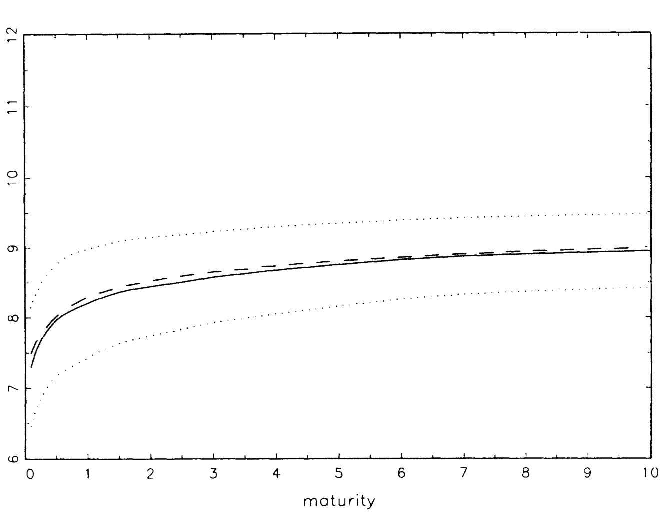

Figure 8. Fit of the Two-Factor Affine Model. The figure shows the average actual and fitted term structures, as well as the root mean squared error of the difference between fitted and observed values, in the three-factor affine model: ———, average yield curve; – –, fitted curve; · · · ·, RMSE; —·· —, bands.

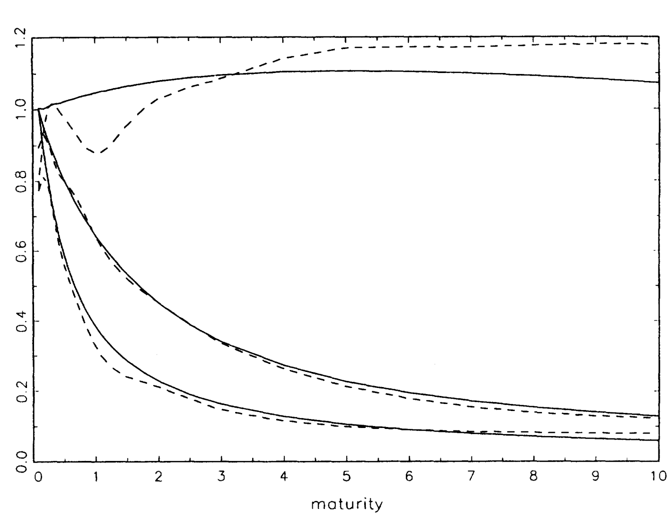

Figure 9. Regression of Observed Yields on Fitted Factors. The figure shows the coefficients of a regression (in first differences) of the observed time series of yields on the time series of fitted factors in the three-factor affine model.

negative correlation of that factor with the short rate (approximately the sum of the first and second factor in our model).

Figure 8 shows the fit of the three-factor models. The average fitted term structure is very close to the observed average term structure. The three-factor model apparently is able to capture the steep short end of the yield curve. The improved fit to the average is also obvious from the increase in the likelihood, although the variance of the residuals is not much lower than in the two-factor model. Another important improvement of the three-factor model over the two-factor model is in the residual serial correlation. The first-order serial correlation is down to around .25, and there is almost no correlation at a 12-month horizon (see Table 10).

As before, I regress the observed yields on the estimated factors. Figure 9 shows the results of such a regression in first differences. On the whole, the estimated coefficients correspond quite well to the model values, although there is some misfit on the first factor for short-term rates. The first factor has a slow mean reversion and low instantaneous volatility, however, and is therefore less important for the volatility of the short-term interest rates than the other two factors.

Table 10. Residuals of the Three-Factor Affine Model

	Maturity			
	.25	1	5	10
Mean	−.0644	−.0647	−.0317	−.0224
Standard deviation	.7705	.7170	.4994	.3981
ρ_1	.2448	.2650	.2496	.1594
ρ_{12}	−.1292	−.0206	.0294	.0209
Correlation matrix				
	1.0000	.9018	.7288	.6327
	.9018	1.0000	.8901	.7779
	.7288	.8901	1.0000	.9218
	.6327	.7779	.9218	1.0000

NOTE: This table reports descriptive statistics (sample mean, standard deviation, and serial correlations, ρ_k) of the residuals of the three-factor affine term-structure model. The scale of the residuals is percentage points.

As a final specification test, I separated the parameters of the cross-section and time series dimensions. This leads to a very large number of parameters, and full estimation of the unrestricted model turned out to be infeasible. Instead, I performed a Lagrange multiplier (LM) test of the restriction that the time series and cross-section parameters are equal. The value of this test is 31.23, which should be compared with the critical values of a chi-squared distribution with 9 df. The 5% critical value is 16.91, so formally I still reject the model specification. The rejection is not as strong as before, however, and in any case one should be cautious with the LR and LM tests because they assume normality of the prediction errors.

From this analysis I conclude that a three-factor model provides an adequate fit of the term-structure data. On average, the model captures the observed shape of the yield curve and the residuals show little persistence.

5. CONCLUSION

In this article, I provide an empirical analysis of the affine class of term-structure models proposed by Duffie and Kan (1996) on a panel of monthly U.S. interest-rate data. I show that a Kalman-filter QML estimator based on the conditional mean and variance has quite good properties. This estimation method fully exploits the panel nature of the data, combines the time series and cross-section information in a theoretically consistent way, and is relatively easy to implement compared with simulation-based estimators.

I estimated affine models with one, two, and three factors. The results show that the one-factor models are misspecified: The fit is not very good, and there is strong residual serial correlation. A formal test of equality of the parameters in the bond-pricing equations (the cross-section dimension of the model) and the factor dynamics (the time series dimension) also rejects the model restrictions. The two-factor model gives a substantial improvement over the one-factor model, but it has some problems fitting the steep initial part of the yield curve. Adding a third factor alleviates most of the problems with the two-factor model. The fit of the average yield curve is quite good and the serial correlation of the residuals is low, indicating that the dynamics of the yield curve are well captured by the model.

The interpretation of the three factors is straightforward. The first factor is highly correlated with the long-maturity interest rates and can therefore be interpreted as the level of the term structure. The mean reversion of this factor is very slow, and it generates almost parallel shifts in the term structure. The second factor has a much stronger mean reversion and affects mainly the short- and medium-term interest rates. The fitted second factor tracks the slope of the term structure, defined as the difference between the three-month interest rate and the 10-year rate, closely. Both the level and the slope factor have time-varying volatilities and significant risk premia and are positively correlated. The third factor is a curvature factor with strong mean reversion and negative correlation with the other two factors.

ACKNOWLEDGMENTS

I thank, without implicating, Michael Brennan, Tom Engsted, Jean-Philippe Lesne, Adrian Pagan, and Peter Schotman, as well as seminar participants at CREST, UCLA, the CEPR European Summer Symposium in Financial Markets 1996 in Gerzensee, the Econometric Society European Meeting 1996 in Istanbul, the European Finance Association meeting 1997 in Vienna, and the American Finance Association meeting 1998 in Chicago, and also an anonymous referee and associate editor for helpful comments.

APPENDIX A: CONDITIONAL MOMENTS OF FACTORS

In this appendix, I show how to derive the conditional mean and variance of the generalized square-root process given in Equation (3), which I repeat here for convenience:

$$dF_t = \Lambda(F_t - \mu)dt + \Sigma(\alpha + \mathcal{B}'F_t)_d^{1/2} dW_t, \qquad (A.1)$$

where $(x)_d$ denotes a diagonal matrix with diagonal elements x_i. If the mean-reversion coefficient matrix is normalized to be diagonal, $\Lambda = \text{diag}(-\kappa_1, \dots, -\kappa_n)$, the stochastic differential equation for F_t can be solved using Ito's lemma,

$$de^{-\Lambda t}(F_t - \mu) = e^{-\Lambda t}\Sigma(\alpha + \mathcal{B}'F_t)_d^{1/2} dW_t, \qquad (A.2)$$

which implies

$$F_{t+h} = \mu + e^{\Lambda h}(F_t - \mu) + \int_0^h e^{\Lambda(h-s)}\Sigma(\alpha + \mathcal{B}'F_{t+s})_d^{1/2}\, dW_{t+s}, \qquad (A.3)$$

where $e^{\Lambda h} = \text{diag}(\exp(-\kappa_1 h), \dots, \exp(-\kappa_n h))$. Because the second part of this sum is a martingale, the conditional mean and variance follow immediately as

$$E(F_{t+h}|F_t) = \mu + e^{\Lambda h}(F_t - \mu) \qquad (A.4)$$

and

$$\text{var}(F_{t+h}|F_t) = \int_0^h e^{\Lambda(h-s)}\Sigma(\alpha + \mathcal{B}'E_t(F_{t+s}))_d \Sigma' e^{\Lambda(h-s)}\, ds. \qquad (A.5)$$

Using that $E_t(F_{t+s}) = \mu + e^{\Lambda t}(F_t - \mu)$ and defining $[\Sigma \text{diag}(\alpha + \mathcal{B}'x)\Sigma']_{ij} = a_{ij} + b_{ij}'x$, I obtain

$$\begin{aligned}\text{var}(F_{t+h}|F_t)_{ij} &= \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)}(a_{ij}' E_t F_{t+s} + b_{ij})\, ds \\ &= \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)}(a_{ij} + b_{ij}'(\mu + e^{\Lambda s}(F_t - \mu)))\, ds \\ &= \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)}(a_{ij} + b_{ij}'\mu)\, ds \\ &\quad + \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)} b_{ij}' e^{\Lambda s}(F_t - \mu)\, ds. \end{aligned} \qquad (A.6)$$

Working out the integrals yields the result

$$\text{var}(F_{t+h}|F_t)_{ij} = \frac{1 - e^{-(\kappa_i+\kappa_j)h}}{\kappa_i + \kappa_j}(a_{ij} + b_{ij}'\mu) + \sum_k \frac{e^{-\kappa_k h} - e^{-(\kappa_i+\kappa_j)h}}{\kappa_i + \kappa_j - \kappa_k} b_{ij,k}(F_{t,k} - \mu_k). \qquad (A.7)$$

This result implies a very simple form for the unconditional variance of F_t,

$$\text{var}(F_t)_{ij} = \frac{a_{ij} + b_{ij}'\mu}{\kappa_i + \kappa_j}. \qquad (A.8)$$

APPENDIX B: THE KALMAN FILTER

All models in this article are estimated using QML based on the following Kalman-filter equations, adapted from Hamilton (1994, chap. 13).

Model:

$$y_t = A + BF_t + e_t, \qquad \text{var}(e_t) = H, \qquad (B.1)$$

and

$$F_t = \Phi F_{t-h} + \eta_t, \qquad \text{var}(\eta_t) = Q_t. \qquad (B.2)$$

Initial Conditions:

$$\hat{F}_0 = E(F_t) \qquad (B.3)$$

and

$$\hat{P}_0 = \text{var}(F_t). \qquad (B.4)$$

Prediction:

$$F_{t|t-h} = \Phi \hat{F}_{t-h} \qquad (B.5)$$

and

$$P_{t|t-h} = \Phi \hat{P}_{t-h} \Phi' + Q_t. \tag{B.6}$$

Likelihood Contributions:

$$u_t = y_t - A - BF_{t|t-h}, \tag{B.7}$$

$$V_t = BP_{t|t-h}B' + H, \tag{B.8}$$

and

$$-2 \ln L_t = \ln |V_t| + u_t' V_t^{-1} u_t. \tag{B.9}$$

Updating:

$$K_t = P_{t|t-h} B' V_t^{-1}, \tag{B.10}$$

$$L_t = I - K_t B, \tag{B.11}$$

$$\hat{F}_t = F_{t|t-h} + K_t u_t, \tag{B.12}$$

and

$$\hat{P}_t = L_t P_{t|t-h}. \tag{B.13}$$

APPENDIX C: RISK PREMIA ON LONG BONDS

In my empirical work I also want to calculate the risk premium on long-maturity bonds. Denote the stochastic process followed by the bond price as

$$dP(\tau) = \mu_{P(\tau)} P(\tau) dt + \sigma_{P(\tau)} P(\tau) dW_t, \tag{C.1}$$

where the dependence of coefficients and prices on time is suppressed. The expected instantaneous return on the bond is the risk-free rate plus a risk premium, which depends on the market prices of risk and the instantaneous standard deviation of the bond return

$$\mu_{P(\tau)} = r + \lambda' \sigma_{P(\tau)}. \tag{C.2}$$

From Ito's lemma, the standard deviation of the bond return is

$$\sigma_{P(\tau)} = -\sigma_F B(\tau), \tag{C.3}$$

where $\sigma_F \sigma_F'$ is the instantaneous variance–covariance matrix of the factors. Given the assumed functional forms for σ_F and λ, I obtain

$$\mu_{P(\tau)} = r - \sum_i \psi_i (\alpha_i + \beta_i' F) B_i(\tau). \tag{C.4}$$

This equation shows that the risk premium on each factor is proportional to the instantaneous variance of that factor, multiplied by the factor loading. If all parameters ψ_i are negative, the risk premia are positive. Because the factor loadings are increasing with maturity, longer bonds will typically have a higher expected return than short bonds.

[Received September 1997. Revised July 1999.]

REFERENCES

Andersen, T., and Lund, J. (1997), "Estimating Continuous-time Stochastic Volatility Models of the Short-term Interest Rate," *Journal of Econometrics*, 77, 343–377.

Babbs, S., and Nowman, K. B. (1999), "Kalman Filtering of Generalized Vasicek Term Structure Models," *Journal of Financial and Quantitative Analysis*, 34, 115–130.

Balduzzi, P., Das, S. R., and Foresi, S. (1995), "The Central Tendency: A Second Factor in Bond Yields," working paper, New York University, Dept. of Finance.

Bams, D., and Schotman, P. (1998), "Direct Estimation of the Risk Neutral Factor Dynamics of Affine Term Structure Models," working paper, University of Maastricht, Dept. of Finance.

Buraschi, A. (1996), "How Large is the Inflation Risk Premium in the U.S. Nominal Term Structure?" IFA Working Paper 232-1996, London Business School, Dept. of Finance.

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

Chen, R., and Scott, L. (1993), "ML Estimation for a Multifactor Equilibrium Model of the Term Structure," *Journal of Fixed Income*, December, 14–31.

Cox, J., Ingersoll, J., and Ross, S. (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53, 385–407.

Dai, Q., and Singleton, K. (1999), "Specification Analysis of Affine Term Structure Models," working paper, Stanford University, Graduate School of Business.

De Jong, F., and Santa-Clara, P. (1999), "The Dynamics of the Forward Interest Curve: A Formulation With State Variables," *Journal of Financial and Quantitative Analysis*, 34, 131–157.

De Munnik, J., and Schotman, P. (1994), "Cross Sectional Versus Time Series Estimation of Term Structure Models: Evidence for the Dutch Bond Market," *Journal of Banking and Finance*, 18, 997–1025.

Duan, J., and Simonato, J. (1995), "Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter," working paper, CIRANO.

Duffie, D., and Kan, R. (1996), "A Yield-factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.

Frachot, A., Lesne, J. P., and Renault, E. (1995), "Indirect Inference Estimation of Factor Models of the Yield Curve," working paper, CREST.

Gallant, A. R., and Tauchen, G. (1996), "Which Moments to Match?" *Econometric Theory*, 12, 657–681.

Geyer, A., and Pichler, S. (1997), "A State-space Approach to Estimate and Test Multi-factor Cox–Ingersoll–Ross Models of the Term Structure," working paper, University of Vienna, Dept. of Statistics.

Gouriéroux, C., Montfort, A., and Renault, E. (1992), "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118.

Hamilton, J. (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press.

Heath, D., Jarrow, R., and Morton, A. (1992), "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60, 77–106.

Hull, J. (1993), *Options, Futures and Other Derivative Securities* (2nd ed.), Englewood Cliffs, NJ: Prentice-Hall.

Jegadeesh, N., and Pennacchi, G. (1996), "The Behaviour of Interest Rates Implied by the Term Structure of Eurodollar Futures," *Journal of Money, Credit and Banking*, 28, 426–446.

Lamoureux, C., and Witte, H. D. (1998), "Empirical Analysis of the Yield Curve: The Information in the Data Viewed Through the Window of Cox, Ingersoll and Ross," working paper, University of Arizona, Dept. of Economics.

Langetieg, T. (1980), "A Multivariate Model of the Term Structure of Interest Rates," *Journal of Finance*, 35, 71–97.

Litterman, R., and Scheinkman, J. (1991), "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, 1, 54–61.

Longstaff, F., and Schwartz, E. (1992), "Interest Rate Volatility and the Term Structure: A Two-factor General Equilibrium Model," *Journal of Finance*, 47, 1259–1282.

Lund, J. (1997), "Econometric Analysis of Continuous-time Arbitrage-free Models of the Term Structure of Interest Rates," working paper, Aarhus

School of Business, Dept. of Finance.

McCulloch, J., and Kwon, H. (1993), "U.S. Term Structure Data, 1947–1991," Working Paper 93-6, Ohio State University, Dept. of Economics.

Pagan, A., and Martin, V. (1996), "Simulation Based Estimation of Some Factor Models in Econometrics," working paper, University of Melbourne, Dept. of Economics.

Pang, K., and Hodges, S. (1996), "Non-negative Affine Yield Models of the Term Structure," working paper, Warwick Business School, Financial Options Research Centre.

Pearson, N., and Sun, T. S. (1994), "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the CIR Model," *Journal of Finance*, 46, 1279–1304.

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

White, H. (1982), "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1–25.