No. 2065

TIME-SERIES AND CROSS-SECTION INFORMATION IN AFFINE TERM STRUCTURE MODELS

Frank de Jong

FINANCIAL ECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

TIME-SERIES AND CROSS-SECTION INFORMATION IN AFFINE TERM STRUCTURE MODELS

Frank de Jong

Discussion Paper No. 2065
February 1999

Centre for Economic Policy Research
90–98 Goswell Rd
London EC1V 7RR
Tel: (44 171) 878 2900
Fax: (44 171) 878 2999
Email: cepr@cepr.org

CEPR Discussion Paper No. 2065

February 1999

ABSTRACT

Time-series and Cross-section Information in Affine Term Structure Models*

In this paper we provide an empirical analysis of the term structure of interest rates using the affine class of term structure models introduced by Duffie and Kan. We estimate these models by combining time-series and cross-section information in a theoretically consistent way. In the estimation we use an exact discretization of the continuous time factor process and allow for a general measurement error structure. We 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. The shocks to these factors are significantly correlated.

JEL Classification: C33, E43
Keywords: term structure, panel data, Kalman filter

Frank de Jong
Department of Financial Management
University of Amsterdam
Roetersstraat 11
1018 WB
Amsterdam
THE NETHERLANDS
Tel: (31 20) 525 5815
Fax: (31 20) 525 5285
Email: fdejong@fee.uva.nl

*The author thanks, without implicating, Michael Brennan, Tom Engsted, Jean-Philippe Lesne, Adrian Pagan, Peter Schotman, and seminar participants at CREST, UCLA, the CEPR European Summer Symposium in Financial Markets, 15 July 1996, Gerzensee, the Econometric Society European Meeting 1996, Istanbul, the European Finance Association meeting 1997, Vienna and the American Finance Association meeting 1998, Chicago for helpful comments.

Submitted 19 January 1999

NON-TECHNICAL SUMMARY

Modelling the dynamics of the term structure of interest rates has been a longstanding challenge for economists. Traditional methods for describing the evolution of interest rates rely on descriptive statistical methods such as principal components and factor analysis. These methods typically identify two or three common factors that drive the term structure. A drawback of these methods is that they do not have a clear underlying economic model for the cross-section of interest rates, i.e. the relation between interest rates in the same period.

Recently, there have been important developments in the theory of interest rates that provide tractable multivariate models describing both the dynamic evolution and the cross-section of interest rates in a theoretically consistent way. The basis of these models is a set of (unobservable) factors that drive the interest rate dynamics and a set of no-arbitrage conditions that specify the cross-sectional relation between bond prices and hence interest rates. Each factor potentially carries a risk premium. A particularly tractable example of these factor based term structure models is the affine model of Duffie and Kan (1997). In that model, all interest rates are affine (linear) functions of the underlying factors. Moreover, the drift and volatility of the interest rates are also affine functions of the level of the factors. This structure makes the model very easy to implement for hedging a portfolio of bonds. The sensitivity with respect to each factor can be seen as a ‘duration’ measure.

In this paper we take this affine model to the data. The paper develops a straightforward estimation method, based on Kalman filters, for the affine model. This method is computationally much less demanding than alternative, simulation based estimators. The method fully exploits the time series and cross-section information in the panel of interest rate data. The panel data approach has the advantage over pure time-series or cross-section methods in that the market prices of factor risk can be estimated. In addition, it yields a natural specification test for the model.

The paper analyses the empirical performance of one, two and three factor implementations of the affine model on a sample of US term structure data. The results show that a model with only one common factor cannot capture the interest rate dynamics. In particular, the short end of the term structure is much more volatile than implied by the fitted model. A two factor model is reasonably well specified. The two factors are closely related to the level and the slope (steepness) of the term structure. Both factors exhibit mean reversion, a level effect in the volatility and are significantly correlated. The market price of risk for both factors is significant and accurately explains the observed risk premia on long and short maturity bonds. The two factor model

cannot account for the volatility hump for medium-range maturities, but adding a third ‘curvature’ factor to the model resolves this problem.

1 Introduction

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. We will 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 analysed separately, but there is a growing literature that estimates term structure models using panel data, i.e. combined cross section and time series data.[1] Typically, the models analyzed are multi-factor versions of the Cox, Ingersoll and Ross (CIR, 1985) model with mutually independent factors. In this paper we 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, Lesne and Renault (1995) estimate a two factor affine model on French data using indirect inference methods. However, they provide little evidence on the fit of the model. Dai and Singleton (1997) estimate a specific three factor affine term structure models on US swap rate data. We extend the analysis of Dai and Singleton in two respects. First, we use a slightly broader cross section of maturities to estimate the affine model. Dai and Singleton use only as many maturities as factors. This is unfortunate because in that case the market price of risk parameters are not always identified. We 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. The second way in which we differ from Dai and Singleton is in the correlation and volatility structure. Dai and Singleton (1997)

[1] An undoubtedly incomplete list is Babbs and Nowman (1997), Bams and Schotman (1997), Buraschi (1996), Chen and Scott (1992), De Jong and Santa-Clara (1998), De Munnik and Schotman (1994), Duan and Simonato (1995), Frachot, Lesne and Renault (1995), Geyer and Pichler (1997), Lund (1994,1997), Pagan and Martin (1996), and Pearson and Sun (1994).

model the third factor as a stochastic volatility. Instead, we choose a fairly general model structure and let the data decide on the best specification. In addition, we explicitly compare the empirical performance of affine models with one, two and three factors and assess the contribution of each additional factor.

In the empirical anaysis we 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 appoach fully exploits the restrictions imposed by the term structure model is therefore expected to give more accurate estimates of the dynamics of the term structure. Secondly, 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 equation).

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. We allow for measurement error on all observed maturities and integrate out the latent factors by the Kalman filter. The prediction errors from the Kalman filter can be used to construct a quasi likelihood function, based on the exact conditional mean and variance of the of the factors. Estimation of the model is then by Quasi Maximum Likelihood, maximizing this quasi likelihood function. In this paper we 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 we show by a Monte Carlo experiment in this paper 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.[2] The affine term structure model lends itself naturally to simula-

[2] Examples are indirect inference, applied to affine term structure models by Frachot, Lesne and Renault (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 (1997) and Pagan and Martin (1997), or Markov chain Monte Carlo methods, used in a Bayesian analysis by

tion based estimators since 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 paper.

The setup of the paper is as follows. Section 2 describes the theoretical model and discusses identification issues. Section 3 discusses the empirical implementation and estimation of the model. Section 4 gives a description of the data and section 5 discusses the empirical results. Section 6 concludes.

2 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

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

where the expectation is taken under the 'risk neutral' probability measure Q. Duffie and Kan (1996) propose a class of endogenous term structure models in which the short rate is an affine function of a number of 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 we also need the stochastic process for the factors under the 'risk neutral' probability measure Q. Duffie and Kan (1996) assume 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

Lamoureux and Witte (1998).

Wiener process under the equivalent martingale measure Q. The stochastic process for the factors under Q is given by

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

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

Duffie and Kan (1996) show 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

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

$$\frac{dB(\tau)}{d\tau} = B_0 + (\Lambda^*)'B(\tau) - 1/2 \sum_i \sum_j B_i(\tau)B_j(\tau)b_{ij} \tag{9b}$$

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

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, Ingersoll and Ross (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 we obtain a two factor CIR model which is observationally equivalent to the Longstaff and Schwartz (1992) model. Jegadeesh and Pennacchi (1996) propose a model where 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. Dai and Singleton (1997) show that, as long as A_0 and α are unrestricted, the mean of the factors under the P measure, μ, is irrelevant for the bond prices. We therefore normalize $\mu = 0$.[3] With this normalization, the expectation of the instantaneous short rate is $\theta \equiv \mathsf{E}^P[r_t] = A_0$, and the vector α can be interpreted as the average volatility of the factors.

Pang and Hodges (1996) show 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 we normalize $B_0 = \iota$, so that the instantaneous interest rate r_t equals a constant plus the sum of the factors ($r_t = \theta + \iota' F_t$). The second implication of this result is that only $\Sigma^{-1}\Lambda\Sigma$ can be identified. Without loss of generality we could therefore assume 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 one. We parameterize $\Lambda = \text{diag}(-\kappa_1, .., -\kappa_n)$. For any parameterization, κ is the vector of minus the eigenvalues of the mean reversion matrix. These eigenvalues are always identified and independent of the normalization.

Dai and Singleton (1997) discuss an additional identification problem that occurs in the multivariate Vasicek model. Because $\mathcal{B} = 0$, the Vasicek model is invariant under unitary rotations of the factors.[4] Essentially this means that only $n(n+1)/2$ elements in $\Sigma^{-1}\Lambda\Sigma$ can be identified. To identify the model we assume, in addition to the previous normalizations, that Σ is upper triangular.

Finally, Dai and Singleton (1997) discuss the implications of the existence conditions of Duffie and Kan (1996) for the affine model. The main result for our model

[3] This identification scheme seems to exclude multifactor extensions of the CIR model. However, the usual CIR parameterization $\alpha = 0$, $A_0 = 0$ and $\mathsf{E}(F_t) = \mu$ is equivalent to the parameterization $\alpha = \mathcal{B}'\mu$, $A_0 = \iota'\mu$ and $\mathsf{E}(F_t) = 0$. Since the dimensions of α and μ are the same, the only effective restriction is the one on A_0. When relaxing this restriction the CIR model becomes equivalent to the affine model. This is equivalent to the model of Pearson and Sun (1994), who add an intercept to the CIR pricing equations.

[4] See Dai and Singleton (1997) for an exact definition.

is the following. When treating the off-diagonal elements of Σ as free paramaters, $\mathcal{B}'\Sigma$ 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 $(\tilde{\beta}_{11}, .., \tilde{\beta}_{nn})$ as free parameters.

To summarize, the identifiable parameters[5] in the affine model are $(\kappa, \alpha, \tilde{\beta}, \Sigma, \psi, \theta)$, where $\tilde{\beta}$ is diagonal and the diagonal elements of Σ equal one. In the Vasicek model there is an additional restricion that Σ is upper triangular.

3 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. However, the choice of maturities to construct the factors is rather arbitrary 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 yields for period t in the vector

$$y_t \equiv \begin{pmatrix} Y_t(\tau_1) \\ \vdots \\ Y_t(\tau_k) \end{pmatrix}$$

Also, define the coefficient matrices

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

[5]The identification of the market price of risk parameters, ψ, will be discussed in the next section.

The state space form of the model is

$$y_t = A + BF_t \tag{10a}$$

$$F_{t+h} = \Phi F_t + \nu_{t+h} \tag{10b}$$

where h is the time interval between two observations.

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: $\mathsf{E}[F_{t+h}|F_t] = \Phi F_t$ and $\mathrm{Var}(\nu_{t+h}) = \mathrm{Var}(F_{t+h}|F_t) \equiv q(F_t)$. Appendix A shows that the conditional mean and variance are exact affine functions of the current level of the factors and depend on the parameters $(\kappa, \alpha, \tilde{\beta}, \Sigma)$.

The first equation in this system is the measurent equation. The coefficients of the measurement equation are functions of the parameters under the risk neutral distribution, $(\mu^*, \Lambda^*, \alpha, \tilde{\beta}, \Sigma, \theta)$. The mean of the factors under the risk neutral distribution, μ^*, is identified from the cross section of observed yields as long as the dimension of y_t is at least $n+1$ (the additional yield is necessary to identify θ). Because we fixed the mean under P of the factors by imposing $\mu = 0$, the parameters ψ are exactly identified from relation (6).[6]

The state space form also highlights that the panel structure imposes overidentifying restrictions on the model, in particular the restrictions on Λ^* given in relation (5), which involve the parameters κ and Σ, and the equality of the parameters $(\alpha, \tilde{\beta})$ 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. The important issue is which assumptions to make on the measurement error structure. Chen and Scott (1993) estimate a model with two factors and four maturities. They assume 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,

[6] Pearson and Sun (1994) and Dai and Singleton (1997) estimate an n factor affine model using n series of zero-coupon yields. In the constant volatility case this means that at least one of the elements of μ^* and hence ψ is not identified. When the volatility depends on the level of the factors, all elements of ψ are identified from the restrictions on Λ^*.

or linear combinations thereof, are assumed to be measured with a normally distributed measurement error. The estimation method is Maximum Likelihood. A number of papers, e.g. Duan and Simonato (1995), Geyer and Pichler (1995), and Jegadeesh and Pennacchi (1996) assume that all interest rates are observed with some measurement error, which is both serially and cross-sectionally uncorrelated. De Jong and Santa-Clara (1998) have a similar econometric approach for estimating a Markovian model of the Heath, Jarrow and Morton (1992) type. Bams and Schotman (1997) follow a similar approach but allow for some correlation between the errors for different maturities. Lund (1994) and Frachot, Lesne and Renault (1995) point out that a diagonal error covariance matrix is not robust under linear transformations of the data. They propose to use a more general, non-diagonal, cross sectional correlation matrix for the measurement errors.

Following these suggestions we assume that the measurement errors have zero mean, 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 papers and, more importantly, 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

$$y_t = A + BF_t + e_t, \qquad Var(e_t) = H \tag{11a}$$

$$F_{t+h} = \mu + \Phi(F_t - \mu) + \nu_{t+h}, \qquad Var(\nu_{t+h}) = q(F_t) \tag{11b}$$

Given this state space setup, the most convenient way to estimate the parameters is by Quasi Maximum Likelihood 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 constistent by the Quasi Maximum Likelihood principle.

There is one subtle problem with this argument which 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.[7] 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).[8] Lamoureux and Witte (1998) propose a Bayesian approach where 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 we 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, 1, 5 and 10 years were caclulated from the model equation (8). Then we 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 $(\theta, \kappa, \alpha, \beta, h)$ are estimated using the Kalman filter QML estimator. In the estimation, the market price of risk parameter is treated as known. The reason for this is that we want to concentrate on the effects of the discretization of the transition equation, in which the market price of risk doesn't play a role.

We 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. In the second experiment, we pick the paramters 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 2,

[7] See Lund (1997) for an extensive discussion of this point.

[8] This is the approach followed by Frachot, Lesne and Renault (1995), Pagan and Martin (1996) and Dai and Singleton (1997).

for 100 replications and for two sample sizes, 300 and 1200 observations. The main conclusion we 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 in Lund (1997), who suggests that, for parameters typically found in estimates of term structure models, the bias in the QML estimator is not particularly large. Therefore, we refrain from using computationally intensive simulation-based estimation techniques and report the QML estimates.

4 Data

Our database is the extended McCulloch dataset (McCulloch and Kwon, 1993). This dataset contains monthly observations on US 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 which were calculated from prices of coupon bonds using McCulloch's interpolation method. From 1985 only bonds which 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 we 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, we picked the maturities for inclusion in the yield vector as follows. For the maturities over 10 years, the bond data is 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. We feel more confident using interest rates with maturities of 3 months and longer. Since we have a full error covariance matrix, the number of parameters to be estimated is potentially large. To keep the estimation feasible we confine ourselves to four maturities: three months, one year, five years and ten years.

Figure 1 graphs the data and Table 1 gives some descriptive statistics. The long maturity interest rates are somewhat 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. Since this is also a period with high levels of interest rates, the data give some intuitive support for models where the conditional variance depends on the level of the interest rates.

5 Empirical results

One of the main issues in term structure modeling is the number of factors to include. In this section we therefore discuss the empirical results on one, two and three factor affine models. For each model we provide an extensive specification analysis. Estimation of all models is by the Kalman filter QML estimator described in section 3. Heteroskedasticity consistent standard errors are calculated by the method of White (1982).

5.1 One factor models

In this section we 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 (9a)–(9b) have analytical solutions for $A(\tau)$ and $B(\tau)$. In the one factor model equations (3) and (4) specialize to[9]

$$dr_t = \kappa(\mu - r_t)dt + \sqrt{\tilde{\alpha} + \beta r_t}dW_t \qquad (12a)$$

$$dr_t = \kappa^*(\mu^* - r_t)dt + \sqrt{\tilde{\alpha} + \beta r_t}dW_t^* \qquad (12b)$$

and the term structure is given by $P_t(\tau) = \exp\left[-\tilde{A}(\tau) - B(\tau)r_t\right]$ where the functions $\tilde{A}(\tau)$ and $B(\tau)$ can be found from a straightforward generalization of the standard CIR equations which are given e.g. in Hull (1993):

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

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

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

[9] Here we follow the usual convention in one factor models to specify the process for the instantaneous short rate. Of course, this is consistent with our general model by defining the factor as $F_t = r_t - \mu$. The relation between $\tilde{\alpha}$ and the original parameter is $\alpha = \tilde{\alpha} + \beta\mu$.

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

$$\begin{aligned} \tilde{A}(\tau) &= R_\infty(\tau - B(\tau)) + \frac{\tilde{\alpha}}{4\kappa} B(\tau)^2 \\ B(\tau) &= \frac{1 - \exp(-\kappa\tau)}{\kappa} \end{aligned}$$

where $R_\infty \equiv \mu - \frac{\psi\tilde{\alpha}}{\kappa} - \frac{\tilde{\alpha}}{2\kappa^2}$ is the yield on infinite maturity bonds.

The parameters to be estimated are the mean reversion coefficient κ, the long run mean of the factor μ, 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 0.06. The implied mean reversion under the risk-neutral distribution is even slower: κ^* in the affine model is 0.0014, which implies a half life of around 500 years.[10] 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, $\alpha = \tilde{\alpha} + \beta\mu$ 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 zero. 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, see e.g. Chan et al. (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.[11] The estimated risk premium for a ten year bond is around 1.65% annually, which corresponds quite well to the observed risk premium.

Given the estimated parameters we can construct residuals of the model, defined as the difference between the observed yields and the predicted yields, and therefore

[10] The half-life of the factor is defined as $\ln(2)/\kappa^*$.
[11] The risk premia are derived in Appendix C.

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 zero 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 0.9 or higher, the 12th order autocorrelations are around 0.5.

The bad fit of the one factor model is illustrated in Figure 2, which graphs the average of the fitted[12] and observed term structures for all maturites from 1 month to 10 years, as well as the standard deviation of the residuals for all maturities. 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 on a constant and the estimated factors. Since the data are close to being non-stationary, this regression is done in first differences. The regression coefficients should be more or less the same as the factor loadings obtained from the term stucture 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.

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 folowing 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 test.[13] 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 esti-

[12] The fitted term structure was calculated from the smoothed estimates of the factors. The smoothing equations of Hamilton (1994, Ch. 13.6) were used to calculate these.

[13] The test statistic is 17.76, which is larger than the 5% critical value of a chi-square distribution with 3 degrees of freedom.

mates. To fit the rather flat shape of the 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 model. For the CIR model, the estimate of κ^* is negative, but the shape of the $B(\tau)$ curve is not very different from the other models for the maturities we consider.[14]

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. We therefore now turn to an analysis of multifactor models.

5.2 Two factor models

In this section we present estimates of affine term structure models with two factors. From the discussion on identification in section 3 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 θ; the variance parameters α and $\tilde{\beta}$; the market prices of risk ψ. In addition to the most general model we estimate some restricited 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 we consider is the generalized Vasicek model with constant volatilities ($\tilde{\beta} = 0$). We 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. Figure 6 graphs 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, $B_1(\tau)$

[14] Strictly speaking, a negative mean reversion coefficient implies an explosive process for the instantaneous interest rate under the risk-neutral distribution, but the functions $A(\tau)$ and $B(\tau)$ are well-defined for such values.

should converge to zero for large τ but apparently this convergence is so slow that it is hardly detectable at horizons up to ten years. So, although the model implies constant infitite 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 we consider.

To interpret the factors, we graph the fitted factors together with functions of the data in Figure 4. The first panel of graphs the first factor (with the estimated long run average θ 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 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 ten year rate. These results are in line with the results of Litterman and Scheinkman (1991), who document a level factor and a slope factor in a pricipal components decomposition of the term structure. Balduzzi, Das and Foresi (1991) model the second factor as a central tendency, which is quite similar to a time-varying slope.

A strong result of our analysis is the significant correlation between the factors, which is evident from the non-zero off-diagonal elements of Σ. Indeed, the Likelihood Ratio 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 $\chi^2(2)$ distribution. To see how this correlation affects the dynamics of the factors we write the model is in feedback form, with a non-diagonal mean reversion matrix Λ and instantaneously uncorrelated factors. This form is obtained by pre-multiplying the factors F_t by Σ^{-1}. The rotated factors $\tilde{F}_t = \Sigma^{-1}F_t$ follow the diffusion (under P)

$$d\tilde{F}_t = \Lambda \tilde{F}_t dt + \begin{pmatrix} \sqrt{\alpha_1 + \tilde{\beta}_{11}\tilde{F}_{1t}} dW_{1t} \\ \vdots \\ \sqrt{\alpha_n + \tilde{\beta}_{nn}\tilde{F}_{nt}} dW_{nt} \end{pmatrix} \tag{13}$$

where $\Lambda = \Sigma^{-1}\text{diag}(-\kappa_1, .., -\kappa_n)\Sigma$. With the numbers from Table 6, the mean reversion matrix in feedback form is

$$\Lambda = \begin{pmatrix} -0.0499 & 0.0590 \\ 0.3375 & -1.2898 \end{pmatrix}$$

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 magnitute of the parameters should be related to the volatility of the factors; the conditional variance of the second factor is about ten 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 zero (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 ten year bond is 1.64%, split over the first factor (0.34%) 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. An inspection of the results reveals that most of the improved fit is for the three month rate. For maturities of one year and longer the residuals are still positive on average, indicating that the model overestimates these interest rates. Figure 5 graphs the average fitted term structure and the standard deviation of the residuals. The figure clearly shows that, 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 6 shows the fitted slope coefficients of a time series regressions 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. However, for the second factor 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, we 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, we 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 $(\kappa, \alpha, \tilde{\beta})$ in the cross section and time series dimension is 92.88, much larger than the 5% critical value of a $\chi^2(6)$ 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.

5.3 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 (1997) argue that a third factor is necessary to fully describe the term structure. Both papers model the third factor as a stochastic volatility factor. Instead, we estimate a general three factor model with an affine volatility structure. Although this model has many parameters (for example, Σ has 6 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 3 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:

$$\Lambda = \begin{pmatrix} 0.0429 & 0.0754 & -0.1293 \\ 1.6079 & -1.2028 & -1.3894 \\ 3.2622 & 0.8283 & -3.8474 \end{pmatrix}$$

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, we 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 1 year interest rate, minus the sum of the three month rate and the 10 year rate.[15] 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 in Dai and Singleton (1997), who find that their stochastic volatility factor is closely related to the curvature of the term structure. They also document a 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 0.25, and there is almost no correlation at a 12 month horizon.

Like before we 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. However, the first factor has a slow mean reversion and low instantaneous volatility and is therefore less important for the volatility of the short term interest rates than the other two factors.

As a final specification test we separated the parameters of the cross section and time series dimension. This leads to a very large number of parameters and full estimation of the unrestricted model turned out to be infeasible. Instead, we performed a Lagrange Multiplier 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-square distribution with 9 degrees of freedom. The 5% critical value is 16.91, so formally we still reject the model specification. However, the rejection is not as strong as before and in any case one should be cautions with the LR and LM tests since they assume normality of the prediction errors.

From this analysis we 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 deviations from the model (the residuals) show little persistence.

[15] This definition of the curvature factor is more direct than the curvature factor proposed by Dai and Singleton (1997), who define curvature as the residuals of a regression of the two year rate on the six month and 10 year rates.

6 Conclusion

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

We 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 3 month interest rate and the 10 year rate, closely. Both the level and the slope factor have time varying volatilities, 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.

A Conditional moments of factors

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

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

The mean reversion coefficient matrix is normalized to be diagonal, $\Lambda = \text{diag}(-\kappa_1, .., -\kappa_n)$. The stochastic differential equation for F_t can be solved using Ito's lemma

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

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

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

$$\mathsf{E}(F_{t+h}|F_t) = \mu + e^{\Lambda h}(F_t - \mu)$$

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

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

$$\begin{aligned}
\text{Var}(F_{t+h}|F_t)_{ij} &= \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)}(a'_{ij}E_tF_{t+s} + b_{ij})ds \\
&= \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)}(a_{ij} + b'_{ij}(\mu + e^{\Lambda s}(F_t - \mu)))ds \\
&= \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)}(a_{ij} + b'_{ij}\mu)ds + \int_0^h e^{-(\kappa_i+\kappa_j)(h-s)} b'_{ij} e^{\Lambda s}(F_t - \mu)ds
\end{aligned}$$

Working out the integrals yields the result

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

Note that this results implies a very simple form for the unconditional variance of F_t,

$$\text{Var}(F_t)_{ij} = \frac{a_{ij} + b'_{ij}\mu}{\kappa_i + \kappa_j}.$$

B The Kalman filter

All models in this paper are estimated using Quasi Maximum Likelihood based on the following Kalman filter equations, adapted from Hamilton (1994, Ch.13).

Model

$$\begin{aligned} y_t &= A + BF_t + e_t, \qquad Var(e_t) = H \\ F_t &= \Phi F_{t-h} + \eta_t, \qquad Var(\eta_t) = Q_t \end{aligned}$$

Initial conditions

$$\begin{aligned} \hat{F}_0 &= \mathsf{E}(F_t) \\ \hat{P}_0 &= \mathrm{Var}(F_t) \end{aligned}$$

Prediction

$$\begin{aligned} F_{t|t-h} &= \Phi \hat{F}_{t-h} \\ P_{t|t-h} &= \Phi \hat{P}_{t-h} \Phi' + Q_t \end{aligned}$$

Likelihood contributions

$$\begin{aligned} u_t &= y_t - a - BF_{t|t-h} \\ V_t &= BP_{t|t-h}B' + H \\ -2 \ln L_t &= \ln |V_t| + u_t' V_t^{-1} u_t \end{aligned}$$

Updating

$$\begin{aligned} K_t &= P_{t|t-h} B' V_t^{-1} \\ L_t &= I - K_t B \\ \hat{F}_t &= F_{t|t-h} + K_t u_t \\ \hat{P}_t &= L_t P_{t|t-h} \end{aligned}$$

C Risk premia on long bonds

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

$$dP(\tau) = \mu_{P(\tau)} P(\tau) dt + \sigma_{P(\tau)} P(\tau) dW_t$$

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

$$\mu_{P(\tau)} = r + \lambda' \sigma_{P(\tau)}$$

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

$$\sigma_{P(\tau)} = -\sigma_F B(\tau)$$

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

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

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. Since the factor loadings are increasing with maturity, longer bonds will typically have a higher expected return than short bonds.

References

ANDERSEN, T. AND J. LUND (1997), Estimating continuous-time stochastic volatility models of the short-term interest rate, *Journal of Econometrics* **77**, 343-377.

BABBS, S., AND K.B. NOWMAN (1997), Kalman filtering of generalized Vasicek term structure models, working paper, First National Bank of Chicago.

BALDUZZI, P., S.R. DAS, AND S. FORESI (1995), The central tendency: A second factor in bond yields, working paper New York University.

BAMS, D., AND P. SCHOTMAN (1997), A panel data analysis of affine term structure models, working paper, University of Maastricht.

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.

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.

CHEN, R., AND L. SCOTT (1993), ML estimation for a multifactor equilibrium model of the term structure, *Journal of Fixed Income*, December.

COX, J., J. INGERSOLL AND S. ROSS (1985), A theory of the term structure of interest rates, *Econometrica* **53**, 385-407.

DAI, Q., AND K. SINGLETON (1997), Specification analysis of affine term structure models, working paper, Stanford University.

DE JONG, F., AND P. SANTA-CLARA (1998), The dynamics of the forward interest curve: A formulation with state variables, working paper.

DE MUNNIK, J. AND P. SCHOTMAN (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 J. SIMONATO (1995), Estimating and testing exponential-affine term structure models by Kalman filter, working paper CIRANO.

DUFFIE, D. AND R. KAN (1996), A yield-factor model of interest rates, *Mathematical Finance* **6**, 379-406.

FRACHOT, A., J.P. LESNE AND E. RENAULT (1995), Indirect inference estimation of factor models of the yield curve, working paper.

GALLANT, A.R. AND G. TAUCHEN (1996), Which moments to match?, *Econometric Theory* **12**, 657-681.

GEYER, A. AND S. PICHLER (1997), A state-space approach to estimate and test multi-factor Cox-Ingersoll-Ross models of the term structure, working paper, University of Vienna.

GOURIÉROUX, C., A. MONTFORT AND E. RENAULT (1992), Indirect inference, *Journal of Applied Econometrics* **8**, S85-118.

HAMILTON, J. (1994), *Time Series Analysis*, Princeton University Press.

HEATH, D., R. JARROW AND A. MORTON (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*, second edition, Prentice-Hall.

HULL, J. AND A. WHITE (1990), Pricing interest rate derivative securities, *Review of Financial Studies* **3**, 573-592.

JEGADEESH, N. AND G. PENNACCHI (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 H.D. WITTE (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.

LANGETIEG, T. (1980), A multivariate model of the term structure of interest rates, *Journal of Finance* **35**, 71-97.

LITTERMAN, R., AND J. SCHEINKMAN (1991), Common factors affecting bond returns, *Journal of Fixed Income* **1**, 54-61.

LONGSTAFF, F. AND E. SCHWARTZ (1992), Interest rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance* **47**, 1259-1282.

LUND, J. (1994,1997), Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rates, working paper, Aarhus School of Business.

MCCULLOCH, J. AND H. KWON (1993), U.S. term structure data, 1947-1991, Ohio State University Working Paper 93-6.

PAGAN, A. AND V. MARTIN (1996), Simulation based estimation of some factor models in econometrics, working paper, University of Melbourne.

PANG, K. AND S. HODGES (1996), Non-negative affine yield models of the term structure, working paper, Financial Options Research Centre, Warwick Business School, UK.

PEARSON, N. AND T.S. SUN (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.

Table 1: Descriptive statistics of the McCullogh 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
minumum	3.38	3.74	5.15	5.72
maximum	16.00	16.35	15.70	15.07

Table 2: Monte Carlo results

This table reports results of a Monte Carlo experiment for the QML estimator of a one factor affine model. The number of simulation runs is 100. In each simulation 300 or 1200 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

300 observations

	true value	median	mean	stdev	t-value
κ	0.0601	0.0601	0.0599	0.0100	-0.2018
θ	6.4642	6.4612	6.4337	0.4617	-0.6606
α	1.0468	1.0528	1.0527	0.0786	0.7495
β	0.3961	0.3906	0.3909	0.0843	-0.6141
h	0.2479	0.2475	0.2478	0.0047	-0.1546

1200 observations

	true value	median	mean	stdev	t-value
κ	0.0601	0.0606	0.0605	0.0037	1.0599
θ	6.4642	6.4755	6.4545	0.1484	-0.6516
α	1.0468	1.0455	1.0458	0.0216	-0.4670
β	0.3961	0.3962	0.3979	0.0284	0.6273
h	0.2479	0.2472	0.2474	0.0026	-1.6976

300 observations

	true value	median	mean	stdev	t-value
κ	1.3056	1.3225	1.3135	0.1083	0.7305
θ	6.1000	6.0903	6.0923	0.0897	-0.8597
α	4.5400	4.5334	4.5825	0.4333	0.9800
β	2.2960	2.3710	2.3311	0.4895	0.7169
h	0.2479	0.2479	0.2481	0.0059	0.4496

1200 observations

	true value	median	mean	stdev	t-value
κ	1.3056	1.3260	1.3296	0.0624	3.8489
θ	6.1000	6.0891	6.0569	0.3337	-1.2920
α	4.5400	4.6095	4.7600	1.9051	1.1548
β	2.2960	2.3796	2.3840	0.5062	1.7382
h	0.2479	0.2481	0.2976	0.4144	1.2006

Table 3: Estimation results one factor affine models

This table reports QML estimates and standard errors for the parameters of the one factor affine, CIR and Vasicek term structure model. For legibility, the estimates of the parameters are scaled as follows: κ, 100μ, $\psi/100$, $10000\tilde{\alpha}$ and 100β. The table also reports the mean reversion parameter κ^*, and the half-life of the factors, $[\ln(2)/\kappa^*]$, under the risk neutral distribution.

	affine	CIR	Vasicek
κ	0.0601	0.0429	0.0222
	(0.0074)	(0.0042)	(0.0028)
$\tilde{\alpha}$	-1.5137		1.9980
	(0.4369)		(0.2284)
β	0.3961	0.2168	
	(0.0637)	(0.0231)	
ψ	-0.1481	-0.1446	-0.0928
	(0.0069)	(0.0110)	(0.0178)
μ	6.4642	5.8099	7.3146
	(0.2465)	(0.5397)	(1.4600)
κ^*	0.0014	0.0116	0.0222
	[495.10]	[60.00]	[31.19]
$2\ln L$	710.45	702.43	677.60

Table 4: Residuals of the one factor affine model

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.

maturity (years)	0.25	1	5	10
mean	-0.7027	-0.2395	0.0760	0.1146
standard deviation	1.5967	1.3009	0.5513	0.3851
ρ_1	0.8957	0.8459	0.4427	0.2031
ρ_{12}	0.4635	0.4265	0.2070	0.1443
correlations				
0.25	1.0000	0.9500	0.6958	0.4123
1	0.9500	1.0000	0.8230	0.5525
5	0.6958	0.8230	1.0000	0.8731
10	0.4123	0.5525	0.8731	1.0000

Table 5: Estimation results for one factor affine models with separate time series parameters

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 ˆ are the time series paramaters. See also the notes at Table 3.

	affine	CIR	Vasicek
κ	0.0230	0.0307	0.0218
	(0.0034)	(0.0049)	(0.0031)
$\tilde{\alpha}$	6.0774		6.0548
	(1.1225)		(0.8864)
β	0.0088	0.5943	
	(0.0002)	(0.0907)	
ψ	-0.0630	-0.0688	-0.0631
	(0.0021)	(0.0068)	(0.0040)
μ	7.4027	7.5788	7.5574
	(0.8220)	(1.0065)	(1.0509)
κ^*	0.0225	-0.0102	0.0218
	[31.39]		[31.70]
$\hat{\kappa}$	0.2183	0.2041	0.2018
	(0.1116)	(0.1242)	(0.1321)
$\hat{\alpha}$	-1.7248		1.9220
	(0.4813)		(0.2185)
$\hat{\beta}$	0.4318	0.1988	
	(0.0788)	(0.0232)	
$2\ln L$	728.21	714.55	693.54

Table 6: Estimation results two factor affine models

This table reports QML estimates and standard errors of the two factor affine and Vasicek term structure models defined in Section 5.2. The estimates are reported as κ_i, σ_{ij}, 100θ, $\psi_i/100$, $10000\alpha_i$ and $100\tilde{\beta}_{ii}$. The table also reports the eigenvalues, κ_i^*, of the mean reversion matrix under the risk neutral distribution, Λ^*, and the associated half-lifes.

	affine	affine uncorr	Vasicek	Vasicek uncorr
κ_1	0.0341	0.0164	0.0234	0.0232
	(0.0074)	(0.0045)	(0.0029)	(0.0029)
κ_2	1.3056	1.1103	0.8424	0.8469
	(0.1126)	(0.0765)	(0.0510)	(0.0498)
α_1	0.51	2.48	1.58	1.52
	(0.28)	(1.00)	(0.20)	(0.19)
α_2	4.54	5.86	6.46	6.51
	(0.71)	(0.68)	(1.28)	(1.17)
$\tilde{\beta}_{11}$	0.3043	0.4261		
	(0.0548)	(0.1357)		
$\tilde{\beta}_{22}$	2.2960	1.8889		
	(0.4908)	(0.3674)		
σ_{12}	0.0470		-0.0118	
	(0.0222)		(0.0260)	
σ_{21}	-0.2688			
	(0.0737)			
ψ_1	-0.0655	-0.0454	-0.0213	-0.0275
	(0.0172)	(0.0016)	(0.0128)	(0.0102)
ψ_2	-0.2061	-0.1449	-0.1378	-0.1558
	(0.0095)	(0.0025)	(0.0167)	(0.0153)
θ	6.10	10.75	11.77	10.80
	(0.88)	(0.17)	(1.34)	(0.82)
κ_1^*	0.0054	-0.0031	0.0234	0.0232
	[127.65]		[29.59]	[29.85]
κ_2^*	0.8411	0.8369	0.8424	0.8469
	[0.82]	[0.83]	[0.82]	[0.82])
$2 \ln L$	1376.40	1333.03	1225.31	1225.11

Table 7: Residuals of the two factor affine model

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.

maturity	0.25	1	5	10
mean	0.0106	0.2331	0.1462	0.1157
standard deviation	0.7462	0.7603	0.5086	0.4042
ρ_1	0.1851	0.4629	0.4178	0.3524
ρ_{12}	-0.1087	0.1782	0.1982	0.1876
correlations				
0.25	1.0000	0.8345	0.6828	0.5990
1	0.8345	1.0000	0.8910	0.7900
5	0.6828	0.8910	1.0000	0.9267
10	0.5990	0.7900	0.9267	1.0000

Table 8: Estimation results two factor affine models with separate time series parameters

This table reports QML estimates and standard errors of the two factor affine and Vasicek term structure models with separate time series coefficients $\hat{\kappa}_i$, $\hat{\alpha}_i$ and $\hat{\beta}_i$. See also the notes at Table 6.

	affine	Vasicek			
κ_1	0.0008	0.0246	$\hat{\kappa}_1$	0.0195	0.0953
	(0.0002)	(0.0030)		(0.0047)	(0.0680)
κ_2	1.6441	0.7520	$\hat{\kappa}_2$	0.9222	0.8944
	(1.0468)	(0.0392)		(0.2296)	(0.4896)
α_1	0.02	0.00	$\hat{\alpha}_1$	3.51	1.79
	(0.01)	(0.00)		(1.50)	(0.22)
α_2	263.81	262.61	$\hat{\alpha}_2$	6.53	6.31
	(160.94)	(30.23)		(6.59)	(1.40)
β_{11}	0.3067		$\hat{\beta}_{11}$	0.2969	
	(0.1642)			(0.1028)	
β_{22}	69.9727		$\hat{\beta}_{22}$	1.8824	
	(50.7893)			(0.3629)	
σ_{12}	0.1450	-0.0288			
	(0.0377)	(0.0284)			
σ_{21}	-0.1804				
	(0.1622)				
ψ_1	-0.0151	-0.0272			
	(0.0111)	(0.0086)			
ψ_2	-0.0162	-0.0110			
	(0.0372)	(0.0020)			
θ	16.08	9.79			
	(8.72)	(0.39)			
κ_1^*	-0.0834	0.0246			
		[28.19]			
κ_2^*	0.5904	0.7520			
	[1.17]	[0.92]			
			$2 \ln L$	1469.28	1299.73

Table 9: Estimation results three factor affine model

This table reports QML estimates and standard errors of the three factor affine term structure model defined in Section 5.3. The estimates are reported as κ_i, σ_{ij}, 100θ, $\psi_i/100$, $10000\alpha_i$ and $100\beta_{ii}$. The table also reports the eigenvalues, κ_i^*, of the mean reversion matrix under the risk neutral distribution, Λ^*, and the associated half-lifes.

κ_1	0.0549 (0.0412)	κ_2	1.7407 (1.1920)	κ_3	3.2117 (0.9569)
α_1	2.48 (6.30)	α_2	18.66 (26.87)	α_3	34.98 (31.51)
β_{11}	0.3498 (0.3151)	β_{22}	3.2055 (2.6778)	β_{33}	0.0013 (0.0129)
σ_{11}	1.0000	σ_{12}	0.0325 (0.0257)	σ_{13}	-0.0460 (0.1376)
σ_{21}	0.3433 (1.7716)	σ_{22}	1.0000	σ_{23}	-0.6806 (0.3892)
σ_{31}	-0.8137 (2.7473)	σ_{32}	-0.3818 (0.6113)	σ_{33}	1.0000
ψ_1	-0.1214 (0.0427)	ψ_2	-0.2179 (0.0668)	ψ_3	-0.2786 (0.0469)
θ	11.27 (14.37)				
κ_1^*	-0.0012	κ_2^*	0.8981 [0.77]	κ_3^*	3.3693 [0.21]
$2\ln L$	1602.96				

Table 10: Residuals of the three factor affine model

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.

maturity	0.25	1	5	10
mean	-0.0644	-0.0647	-0.0317	-0.0224
standard deviation	0.7705	0.7170	0.4994	0.3981
ρ_1	0.2448	0.2650	0.2496	0.1594
ρ_{12}	-0.1292	-0.0206	0.0294	0.0209
correlations				
0.25	1.0000	0.9018	0.7288	0.6327
1	0.9018	1.0000	0.8901	0.7779
5	0.7288	0.8901	1.0000	0.9218
10	0.6327	0.7779	0.9218	1.0000

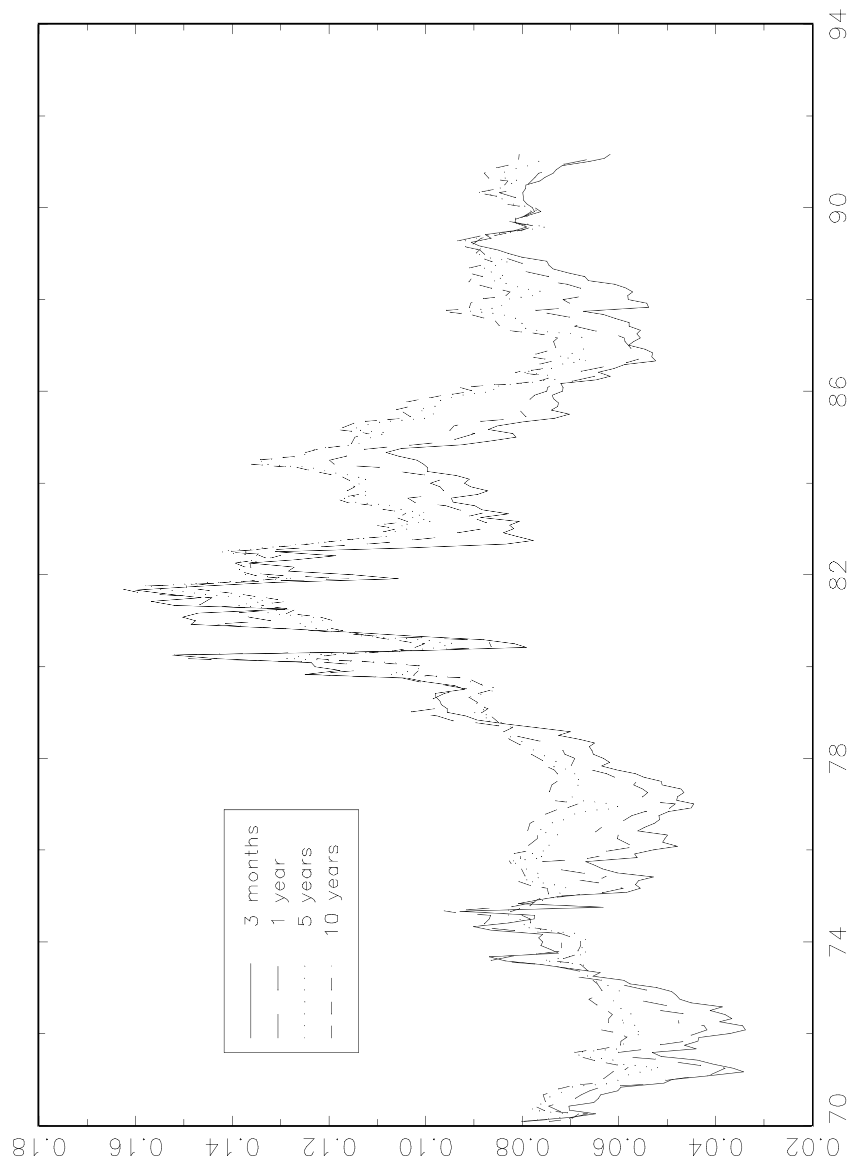

Figure 1: US term structure data.

The figure shows the 3 month, 1, 5 and 10 year zero coupon yields constructed by McCullogh and Kwon (1993) from US treasury coupon bonds.

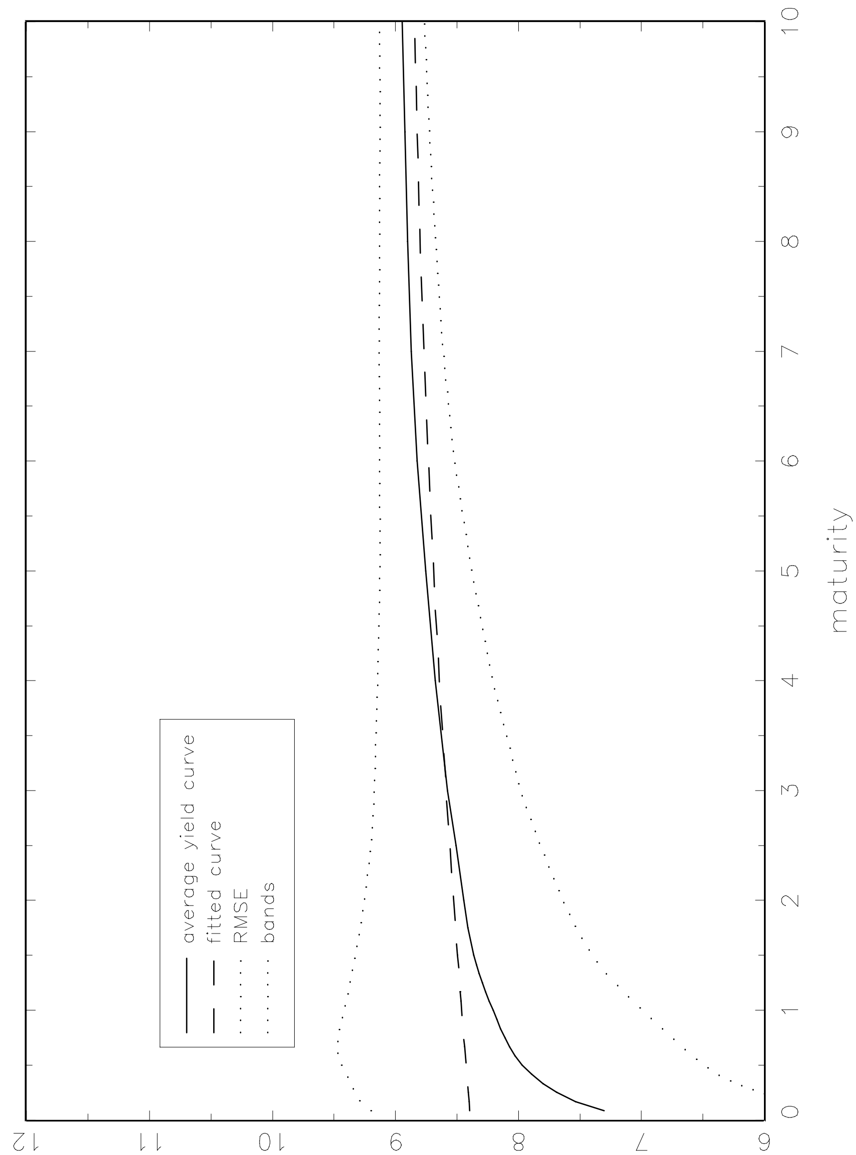

Figure 2: Fit of the one factor affine model.

The figure shows the average actual and fitted term structures, as well as the standard deviation of the residuals, in the one factor affine model.

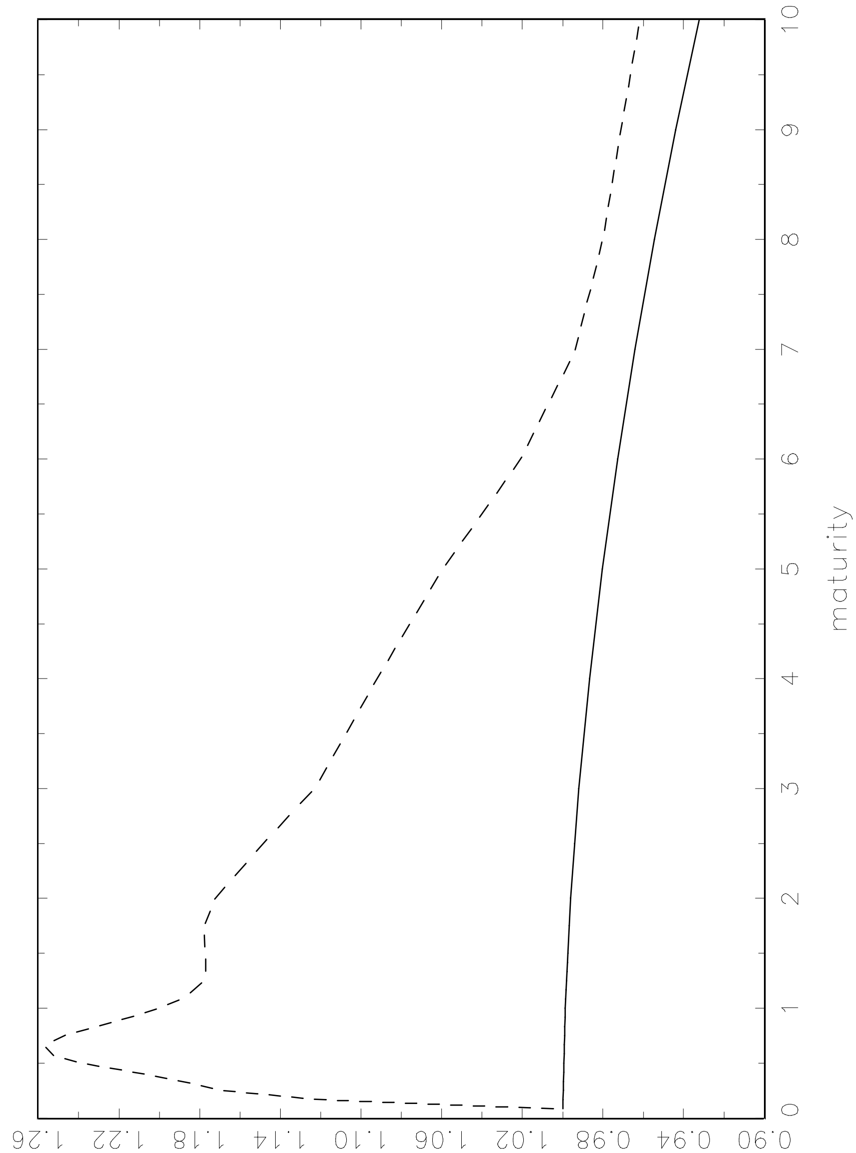

Figure 3: Regression of observed yields on fitted factors.

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.

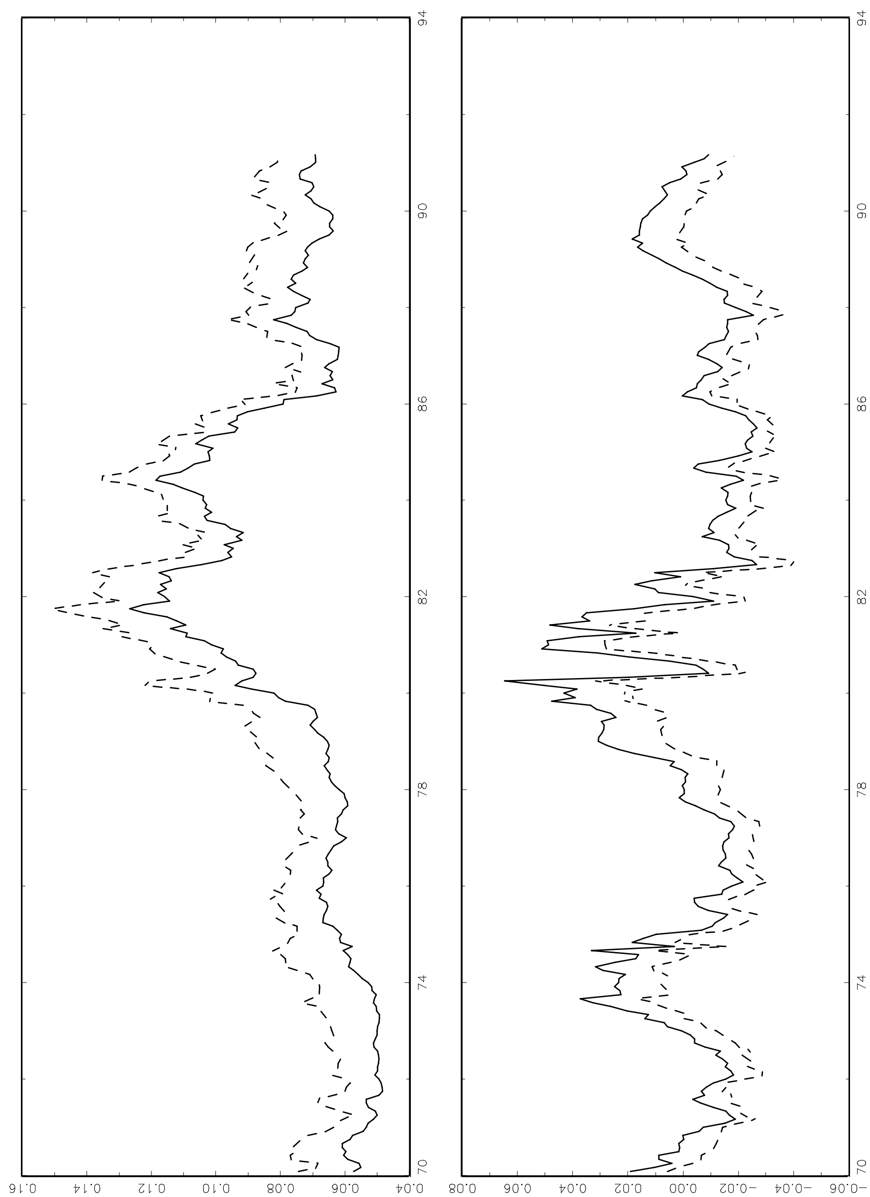

Figure 4: Fitted factors in the two factor affine model.

The figure shows the estimated factors in the two factor affine model.

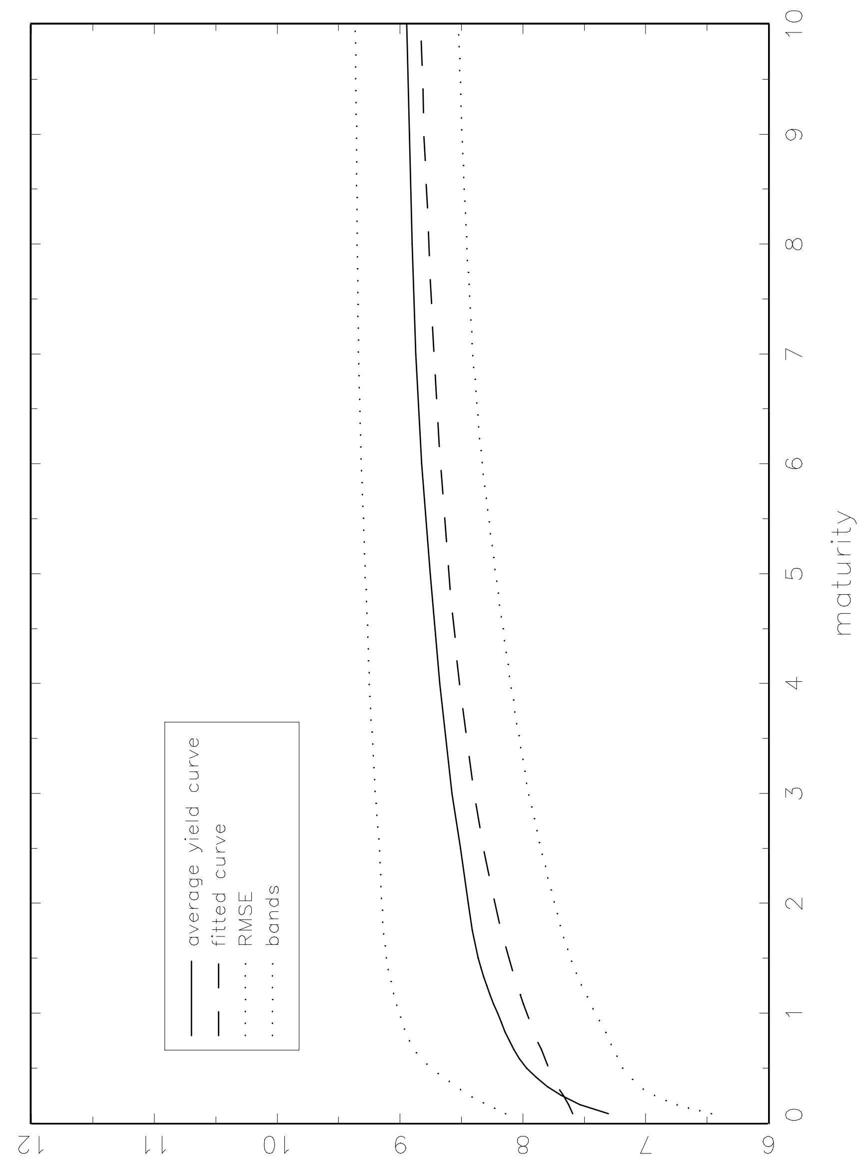

Figure 5: Fit of the two factor affine model.

The figure shows the average actual and fitted term structures, as well as the standard deviation of the residuals, in the two factor affine model.

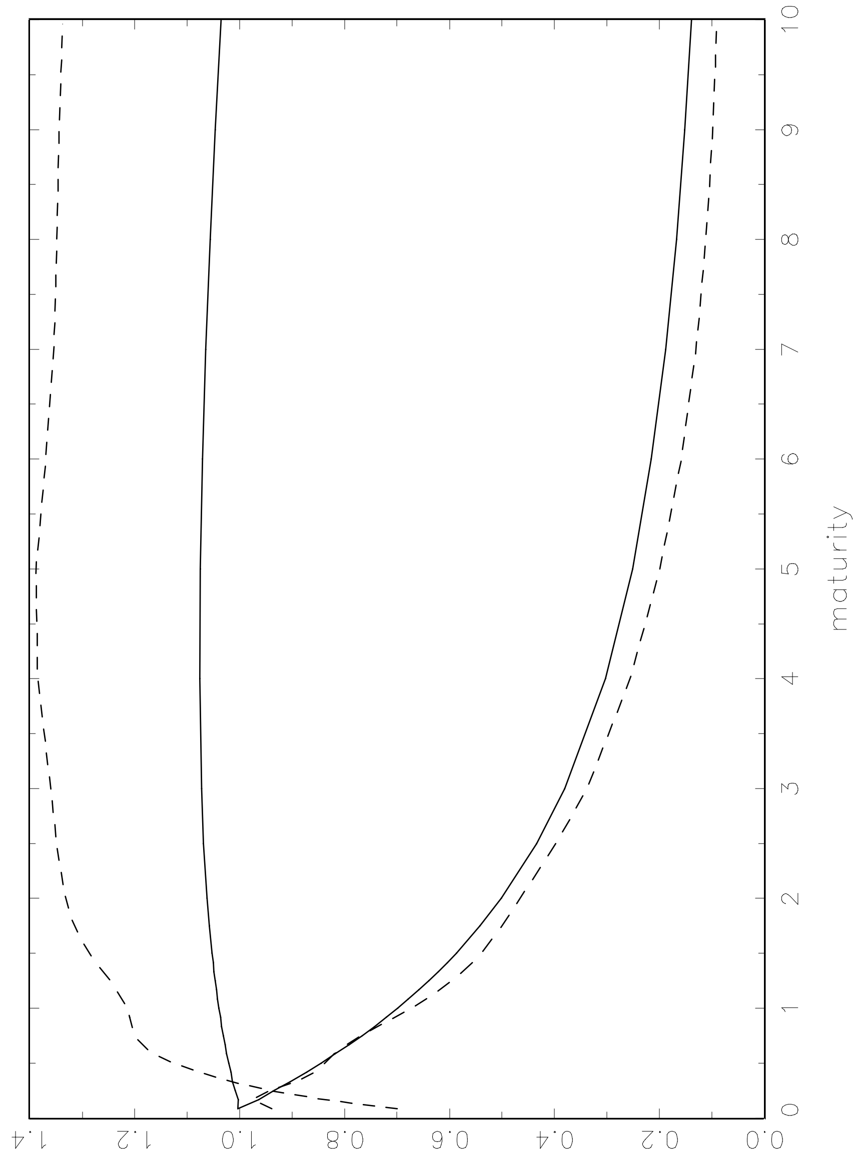

Figure 6: Regression of observed yields on fitted factors.

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 two factor affine model.

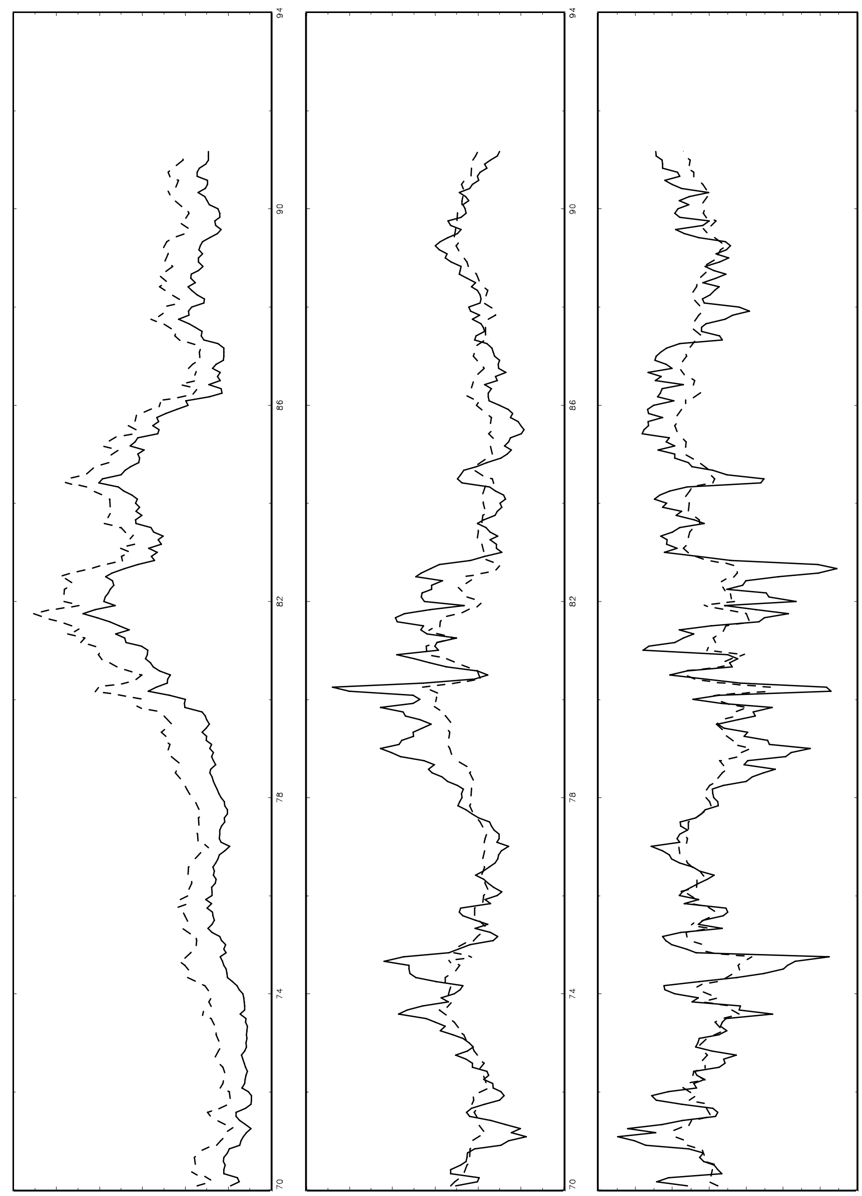

Figure 7: Fitted factors in the three factor affine model.

The figure shows the estimated factors in the three factor affine model.

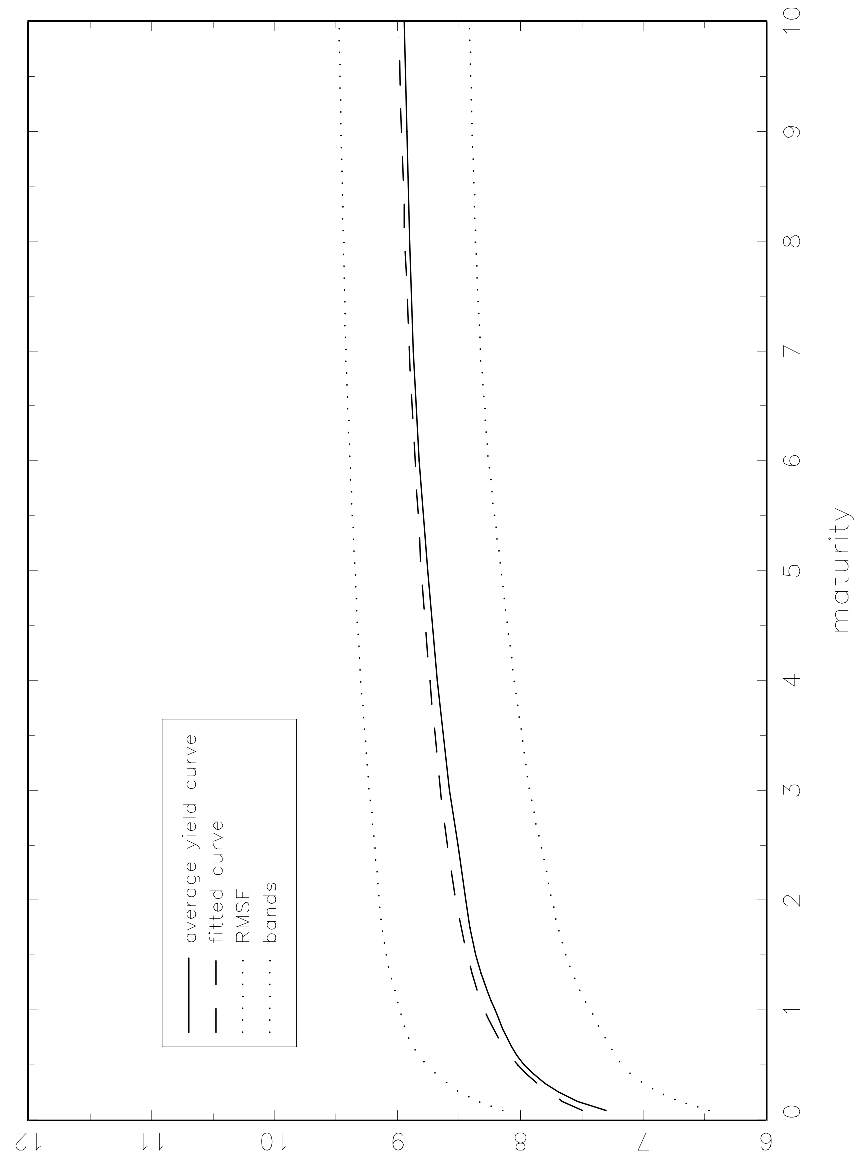

Figure 8: Fit of the two factor affine model.

The figure shows the average actual and fitted term structures, as well as the standard deviation of the residuals, in the three factor affine model.

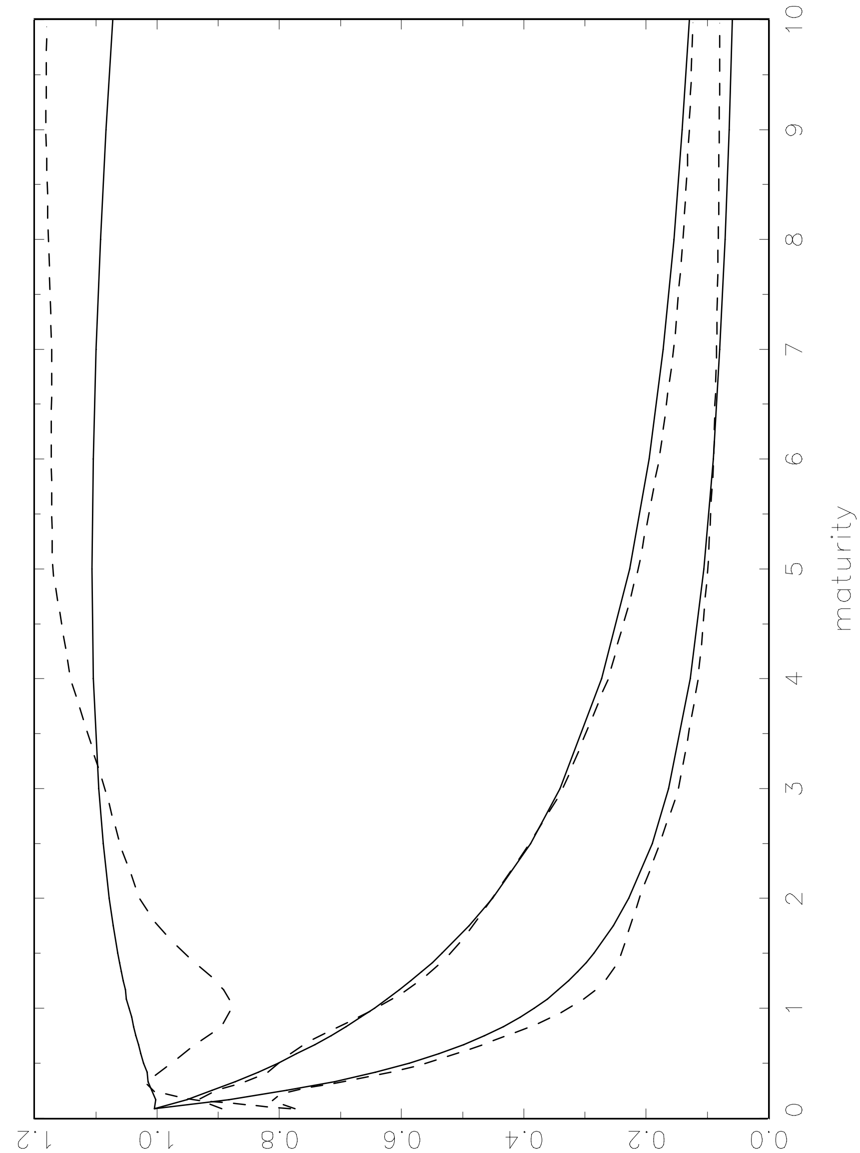

Figure 9: Regression of observed yields on fitted factors.

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.