An empirical analysis of affine term structure models using the generalized method of moments*

Peter de Goeij
K.U. Leuven

Marno Verbeek
K.U. Leuven and Erasmus University Rotterdam

December 22, 2000

Abstract

In this paper we formulate two tractable two-factor affine term structure models, imposing weak assumptions on the distributions of the measurement errors involved in the different yields. Exploiting the implied moment conditions, the models are estimated by the generalized method of moments using weekly term structure data for Germany, Japan, the UK and the USA. Despite our relatively weak assumptions, for each of these countries the overidentifying restrictions tests indicate that the estimated two-factor models should be rejected. Apparently, the fact that many affine term structure models are rejected empirically, is unlikely to be due to the assumptions about the joint distribution of the measurement errors, but more likely to the lack of flexibility to explain certain aspects of the term structure.

JEL-classification: E43, G12

Keywords: Term structure of interest rates, Generalized Method of Moment estimation, Affine term structure models.

*Adress for correspondence: Center for Economic Studies, Naamsestraat 69, B-3000 Leuven, Belgium. tel: +32-16-326650 fax: +32-16-326796 email: Peter.deGoeij@econ.kuleuven.ac.be. We would like to thank Wessel Marquering, Hans Dewachter and Konstantijn Maes for helpful comments and advice. Financial support by the Fund for Scientific Research - Flanders (project FWO G.0220.00) is gratefully acknowledged.

1 Introduction

The challenge for an empirical model of the term structure of interest rates is to jointly describe the cross-sectional shape of the yield curve with its development over time. A range of alternative theoretical term structure models, based upon no arbitrage or equilibrium arguments, has been proposed in the literature; see, for example, Vasicek (1977), Dothan (1978), Langetieg (1980) and Cox, Ingersoll and Ross (1985, CIR hereafter). Most of these models are formulated in continuous time and are tightly parametrized.

A wide range of empirical studies exists examining the term structure of interest rates exploiting either the cross-sectional or the time-series dimension of the data. Some recent examples are Brown and Dybvig (1986), Chan et al. (1992), Brown and Schaefer (1994), Aït-Sahalia (1996) and Gray (1996). Several recent studies focus upon both dimensions of the term structure simultaneously and provide empirical results for affine term structure models; see for example Longstaff and Schwartz (1992), Chen and Scott (1993), Pearson and Sun (1994), Jegadeesh and Pennacchi (1996), Bams and Schotman (1997), Duan and Simonato (1998) de Jong and Santa-Clara (1999), de Jong (2000), Duarte (2000) and Piazzesi (2000). These studies use panel data to estimate the models' parameters more efficiently than is possible from considering a single dimension only. They differ in their assumptions about the number of factors that drive the term structure, their stochastic processes, and the random measurement errors upon (some of) the observed yields. A significant number of studies allow for normally distributed measurement errors upon all maturities and use maximum likelihood combined with a Kalman filter to estimate the model (and to extract the unobserved factors).

In this paper we formulate a discrete time affine term structure model based upon two underlying factors thereby focussing on both dimensions of the term structure simultaneously. Starting from a discrete time pricing kernel approach, assumptions are made about the conditional distribution of the (log) stochastic discount factor as a function of two unobserved factors; see also Backus and Zin (1994), Campbell, Lo and MacKinlay (1997) and Gong and Remolona (1997). From these assumptions direct implications can be derived for the observed term structure data that correspond to an affine yield specification. Next, we allow for additional noise in the observed yields by introducing mean zero measurement errors. In contrast to other studies, we only impose weak distributional assumptions and estimate the model parameters using the generalized method of moments exploiting a fairly small number of moment conditions. This procedure allows us to examine the empirical performance of the models in a way that does not depend upon, for example, assumed joint normality or the imposition of zero measurement errors upon some of the yields. Moreover, the use of GMM enables us to calculate the overidentifying restrictions test that automatically provides a general misspeci…cation test for the a¢ne term structure models estimated in this paper. Such a test is typically not performed in previous studies.

Empirical results for the USA generally indicate that a¢ne term structure models are not very successful in explaining the term structure of interest rates (see, for example, Duan and Simonato, 1998, and Geyer and Pichler, 1998), while empirical studies that investigate the implications of a¢ne term structure models for other countries are scarce. To analyze whether a¢ne term structure models produce more favorable results in other countries, we estimate identical models for four di¤erent countries: the USA, Germany, Japan and the UK. To do so, we use weekly data on zero-coupon bond yields as implied by LIBOR and swap rates.

The remainder of the paper is organized as follows. Section 2 introduces the two-factor a¢ne term structure models that are used in this paper. Section 3 describes the econometric model and the estimation technique. In section 4 we go into more detail about the data we use throughout the paper, while in Section 5 the empirical results are presented. In Section 6 we provide some further explanations of the empirical results and …nally, Section 7 concludes.

2 A Two-Factor A¢ne Model for the Term Structure

A¢ne term structure models have the property that all bond yields are a¢ne[1] functions of a set of state variables (Du¢e, 1992, and Du¢e and Kan, 1996). In the empirical literature, a¢ne models are very popular because of their tractability. Moreover, the class of a¢ne model nests several well-known equilibrium term structure models, like those proposed by Vasicek (1977) and CIR (1985). On the other hand, the assumption in a¢ne yield models that the mean and variance functions are a¢ne in the state variables imposes substantive restrictions upon the shape of the mean yield curve and the term structure of volatility. It is therefore an interesting question to analyze to what extent a relatively simple two-factor a¢ne model is able to capture the observed term structure of interest rates for a number of countries.

[1] A function f is called a¢ne in x if it can be written as $f(x) = ® + ¯x$; for some constants ® and ¯.

It is well known that any asset pricing model can be written as

$$P_t = E_t\{M_{t+1}X_{t+1}\}, \qquad (1)$$

where P_t is the price of an asset at time t, X_{t+1} is the asset's payoff at time t + 1; E_t is the expectation operator conditional upon all information available at time t, and M_{t+1} is the stochastic discount factor (SDF). The stochastic discount factor is a strictly positive random variable that does not vary across assets. Alternative asset pricing models imply different expressions for M_{t+1} (see, e.g., Cochrane, 1999 or Campbell, 2000).

The pricing equation (1) prices all assets, including bonds. Let us denote the price at time t of a zero-coupon bond that pays off one dollar at time t + n by P_{nt}. Then, relationship (1) implies

$$P_{nt} = E_t\{M_{t+1}P_{n-1,t+1}\}. \qquad (2)$$

The price of a single-period bond (n = 1) is

$$P_{1t} = E_t\{M_{t+1}\} = \frac{1}{1 + R^f_{t+1}} \qquad (3)$$

where R^f_{t+1} is the short-term risk free rate. This shows that the conditional expectation of the stochastic discount factor is simply the price of the short-term riskless asset. Solving (2) we obtain

$$P_{nt} = E_t\{M_{t+1}M_{t+2}\ldots M_{t+n}\} = E_t\{M^{(n)}_{t+1}\}, \qquad (4)$$

where $M^{(n)}_{t+1}$ denotes the n-period SDF, defined as the product of n successive one-period SDFs. Thus, the price of an n-period bond is simply the expectation of the n-period SDF, so that for an investor with a horizon of n periods, the n-period bond is riskless (Campbell, 2000). The result in (4) makes clear that a model of bond prices is equivalent to a time-series model of the SDF; see Backus and Zin (1994), who explore this fact empirically. The yield on an n-period bond, denoted by Y_{nt}, is defined as the discount rate that solves

$$P_{nt} = \frac{1}{(1 + Y_{nt})^n}. \qquad (5)$$

The yield equates the bond price to the discounted value of its future payment. Taking logs of (5) results in

$$y_{nt} = -\frac{1}{n}p_{nt}, \qquad (6)$$

where $y_{nt} = \ln(1 + Y_{nt})$ and $p_{nt} = \ln(P_{nt})$. This equality shows that the log bond yield is a simple linear transformation of the log bond price.

Following Campbell, Lo and MacKinlay (1997, p. 428) and Gong and Remolona (1997), we make the assumption that the stochastic discount factor and bond prices are conditionally jointly lognormal. Then (2) can be written as

$$p_{nt} = E_t\{m_{t+1} + p_{n-1,t+1}\} + \frac{1}{2}Var_t\{m_{t+1} + p_{n-1,t+1}\} \tag{7}$$

where $m_{t+1} = \ln(M_{t+1})$. The second term in this expression is due to Jensen's inequality. Affine yield models have the property that all log bond prices are linear in the set of state variables describing the movement of the SDF. That is, the left-hand side of (7) is a linear function of the (potentially unobserved) state variables. Duffie and Kan (1996) derive necessary conditions for the SDF to imply an affine yield model.

The time-series process of the log SDF m_{t+1} can be represented by the sum of its conditional expectation and an innovation term, i.e.

$$m_{t+1} = E_t\{m_{t+1}\} - \varepsilon^*_{t+1},$$

where $E_t\{\varepsilon^*_{t+1}\} = 0$ by construction. The model we investigate is a discrete-time version of the multifactor CIR model, as introduced by Langetieg (1980). Several related empirical studies, e.g., Chen and Scott (1993), Duan and Simonato (1998) and de Jong and Santa Clara (1999) analyze multifactor CIR models for the USA and find that a relatively simple two-factor model is able to describe the average yield curve reasonably well. For example, Chen and Scott (1993) conclude that their two factor model significantly outperforms a (non-nested) three factor model. In line with these studies, we restrict ourselves to two factors and consider a discrete time version of the two-factor model of Longstaff and Schwarz (1992). Consequently, it is assumed that $E_t\{m_{t+1}\}$ is the sum of two independent state variables x_{1t} and x_{2t}

$$E_t\{m_{t+1}\} = -x_{1t} - x_{2t}, \tag{8}$$

where both state variables are assumed to follow a square root process given by

$$x_{i,t+1} = (1 - \phi_i)\mu_i + \phi_i x_{it} + \sigma_i\sqrt{x_{it}}\,\xi_{i,t+1}, \quad i = 1, 2; \tag{9}$$

and where it is assumed that the shocks $\xi_{1,t+1}$ and $\xi_{2,t+1}$ are conditionally normally distributed and mutually independent with mean zero and unit variance. Furthermore, we assume that the error term ε^*_{t+1} is potentially correlated with both the state variables in the following implicit way:

$$\varepsilon^*_{t+1} = \lambda_1\sigma_1\sqrt{x_{1t}}\,\xi_{1,t+1} + \lambda_2\sigma_2\sqrt{x_{2t}}\,\xi_{2,t+1}. \tag{10}$$

The coefficient λ_i, $i = 1, 2$, measuring the correlation between the i-th state variable and the (log)stochastic discount factor, can be interpreted as the market price of risk of the i-th state variable.

The two-factor model proposed by Campbell, Lo and MacKinlay (1997, p. 438-441) is a special case of the above model, obtained by setting λ_2 to zero. In the empirical part of this paper estimation results for the two-factor CIR model will be presented, as well as for the restricted model with $\lambda_2 = 0$ (referred to as the CLM model).

The model presented above meets all conditions on the log SDF and the state variables to result in an affine yield model (see Duffie and Kan, 1993). This implies that the price function for an n-period bond is affine in the two state variables. That is[2]

$$-p_{nt} = A_n + B_{1n} x_{1t} + B_{2n} x_{2t} \tag{11}$$

for some constants A_n, B_{1n} and B_{2n}. Recursive expressions for A_n, B_{1n} and B_{2n} are obtained in a relatively straightforward fashion. Start with $n = 1$ in (7). Because $p_{0,t+1} = 0$, the price of a one-period bond is obtained as

$$\begin{aligned} p_{1t} &= E_t\{m_{t+1}\} + \frac{1}{2}Var_t\{m_{t+1}\} = -x_{1t} - x_{2t} + \frac{1}{2}x_{1t}\lambda_1^2\sigma_1^2 + \frac{1}{2}x_{2t}\lambda_2^2\sigma_2^2 = \\ &= -(1 - \frac{1}{2}\lambda_1^2\sigma_1^2)x_{1t} - (1 - \frac{1}{2}\lambda_2^2\sigma_2^2)x_{2t}. \end{aligned} \tag{12}$$

From this, it follows that $A_1 = 0$, $B_{11} = 1 - \frac{1}{2}\lambda_1^2\sigma_1^2$, and $B_{21} = 1 - \frac{1}{2}\lambda_2^2\sigma_2^2$. Then it is straightforward to show (see Appendix A) that A_n, B_{1n} and B_{2n} obey the following recursive patterns:

$$\begin{aligned} B_{1n} &= 1 + \phi_1 B_{1,n-1} - \frac{1}{2}(\lambda_1 + B_{1,n-1})^2\sigma_1^2; \\ B_{2n} &= 1 + \phi_2 B_{2,n-1} - \frac{1}{2}(\lambda_2 + B_{2,n-1})^2\sigma_2^2; \\ A_n &= A_{n-1} + (1 - \phi_1)\mu_1 B_{1,n-1} + (1 - \phi_2)\mu_2 B_{2,n-1}; \end{aligned} \tag{13}$$

where ϕ_1, ϕ_2, μ_1, μ_2, σ_1, σ_2, λ_1 and λ_2 are the unknown parameters that have to be estimated. Combining (6) and (11) results in an expression for the (log) bond yields that is used in the empirical part of the paper:

$$\begin{aligned} y_{nt} &= \frac{1}{n}A_n + \frac{1}{n}B_{1n}x_{1t} + \frac{1}{n}B_{2n}x_{2t} \\ &= a_n + b_{1n}x_{1t} + b_{2n}x_{2t}. \end{aligned} \tag{14}$$

[2] The minus sign is for convenience only.

While the state variables x_{1t} and x_{2t} are typically unobserved, equation (14) implies severe restrictions on the relationships between different yields at any given point in time, as well as severe restrictions upon the dynamic evolution of the entire yield curve. On the other hand, (14) is sufficiently general to allow for various shapes of the yield curve, depending upon the two underlying factors. In the empirical section we shall exploit some of the implications of (14) to estimate the structural parameters in the model, allowing for deviations from the strict equality in (14) by introducing measurement errors.

3 Econometric Specification

In this section we develop the econometric specification of the model. First, a more general notation for the affine term structure model is introduced. Next, it is shown how the model is extended to incorporate measurement errors and finally it is discussed how the generalized method of moments can be used to estimate the parameters of the model and to test its overall validity.

Assume that m yields are observed, each with a different time to maturity of n_i months, $i = 1, \ldots, m$, and collect these yields in a vector

$$y_t = \begin{bmatrix} y_{n_1 t} \\ \vdots \\ y_{n_m t} \end{bmatrix}. \tag{15}$$

Then define the coefficient vector A and matrix B as

$$A = \begin{bmatrix} a_{n_1} \\ \vdots \\ a_{n_m} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{1n_1} & b_{2n_1} \\ \vdots & \vdots \\ b_{1n_m} & b_{2n_m} \end{bmatrix} \tag{16}$$

so that the affine yield model can be written as

$$y_t = A + Bx_t, \tag{17}$$

where $x_t = (\; x_{1t} \;\; x_{2t} \;)'$ and the underlying stochastic process for the factors is represented by (9) for $i = 1, 2$. As explained in the previous section, the coefficients in A and B are functions of the parameters of the stochastic processes of the factors. Following Longstaff and Schwartz (1992), Chen and Scott (1993) and several others, this paper takes a panel data approach in the sense that the model parameters will be estimated using yield data for a large number of periods.

3.1 Introducing Measurement Errors

The affine model in (17) predicts an exact relation between the factors and the yields. If m exceeds 2, as it will in our empirical section, the strict model in (17) is very unlikely to fit the data. Therefore, we extend the model by introducing random error terms. While these terms capture any kind of specification and measurement error, we shall refer to them as measurement errors, as is common in the literature. Instead of (17), we thus consider the relationship

$$y_t^{obs} = A + Bx_t + v_t; \qquad (18)$$

where y_t^{obs} is the vector of observed yields and $v_t = y_t - y_t^{obs}$ is the vector of measurement errors. An obvious first assumption is that $E\{v_t\} = 0$ but a wide range of alternative additional assumptions can be imposed upon the measurement error terms.

If data are used on as many maturities as factors (in our case two), one can simply assume there is no measurement error at all and follow a procedure that has been suggested by Pearson and Sun (1994). This involves "inverting" the relationship (17) to derive the parameters of the underling stochastic process for the factors (9), $i = 1; 2$. The obvious drawback of this approach is that the results can differ substantially depending upon the maturities that are used, while information contained in any of the other maturities is completely ignored.

Chen and Scott (1993) estimate several affine yield models using data for only four different maturities, assuming that the yields of one, two or three maturities, respectively, are observed without error. The other (linear combinations of) yields are assumed to be measured with a normally distributed measurement error. Under these assumptions, it is possible to apply the "inverting" technique to the maturities observed without error. Obviously, the results will depend upon the choice of maturities that are assumed to be observed without error.

Many other studies assume that all interest rates are observed with some measurement error. For example, Jegadeesh and Pennacchi (1996) and Duan and Simonato (1998) assume that the measurement errors are both serially and cross-sectionally uncorrelated. Bams and Schotman (1997) allow for correlation between the errors for different maturities, where the strength of this correlation depends on the difference in maturity. Lund (1994) and de Jong (2000) assume that the measurement errors are serially uncorrelated, but cross-sectionally correlated with a time-invariant covariance matrix. Because all these studies use maximum likelihood and the Kalman filter to obtain estimates, the common assumption is that the measurement errors are normally distributed.

In contrast to the studies mentioned above, we make extremely weak assumptions in this paper about the measurement errors. The advantages of this are that the parameter estimates are robust and that we can test to what extent the two-factor model is able to explain the observed term structure data even if one allows for very general forms of measurement errors. The potential price for these weak assumptions is that very little structure is imposed upon the model such that parameter estimates may become relatively inaccurate. In the sequel, we shall exploit a number of moment conditions implied by the model to estimate the unknown coefficients and use the overidentifying restrictions test to test the overall specification of the model. Note that with maximum likelihood estimation general misspecification tests (which do not involve a well-specified alternative hypothesis) are typically not performed.

The assumptions we make about the measurement errors v_t can be summarized as follows:

$$E\{v_t\} = 0; \; Var\{v_{it}\} = \omega_i^2; \; i = 1, 2, \ldots, m; \tag{19}$$

$$cov\{v_{1t}, v_{1,t-k}\} = 0; \; k = 1, 2, 3, \ldots \tag{20}$$

That is, all measurement errors have an unconditional mean of zero, while the measurement error on the one-month yield exhibits no autocorrelation. The latter is an identifying assumption to ensure that the two factors drive the dynamics of the one-month interest rate (rather than its measurement error). No restrictions are imposed upon the contemporaneous covariance matrix of all measurement errors, while arbitrary forms of autocorrelation in the measurement errors are allowed for all maturities exceeding one month (up to a finite lag length). The usual assumption is imposed that the measurement errors are independent of the true yields.

3.2 The Estimation Method: GMM

The assumptions in (19), (20) are combined with (18) to derive a number of moment conditions that are exploited to estimate the unknown parameters. In the empirical section, we shall use a set of unconditional moments based upon the expected yields, the variances of the yields and the autocovariances of the one month yield, using weekly observations on a selected number of yields with maturities of one month up to 60 months. The derivations of the appropriate expressions are presented in Appendix B. These unconditional moments are subsequently used in a GMM estimation procedure, as developed by Hansen (1982).

Let us denote the K-dimensional vector of unknown parameters by θ and the R-dimensional vector of moment conditions as $E\{u_t(\theta)\} = 0$, where u_t is a vector function of θ that depends upon observable data (see Appendix B). Furthermore, denote the corresponding vector of sample moments by $g_T(\theta) = T^{-1}\sum_s u_s(\theta)$, where s indexes weeks. Then a GMM estimator for θ is obtained by minimizing the quadratic form

$$g_T(\theta)' W_T g_T(\theta), \tag{21}$$

where W_T is a weighting matrix satisfying plim $W_T = W$, where W is a positive definite symmetric matrix. Different weighting matrices lead to different consistent estimators for θ. Under a number of regularity conditions (Hansen, 1982), the resulting estimator, $\hat{\theta}$ say, is consistent and asymptotically normally distributed with covariance matrix (see Cochrane, 1996)

$$V\{\hat{\theta}\} = \frac{1}{T}(D'WD)^{-1}D'WSWD(D'WD)^{-1}, \tag{22}$$

where D is the $K \times R$ matrix of derivatives of the moments, i.e.

$$D = \text{plim}\,\frac{\partial g_T(\theta)}{\partial \theta}, \tag{23}$$

and S is the covariance matrix of the sample moments. If W is chosen equal to S^{-1}, the GMM estimator is "optimal" in the sense that its asymptotic covariance matrix is as small as possible, and the expression for the covariance matrix reduces to

$$V\{\hat{\theta}\} = \frac{1}{T}(D'S^{-1}D)^{-1}. \tag{24}$$

Below we shall consider two different choices for the weighting matrix. First, we consider a one-step estimator based upon $W_T = I$, the identity matrix, and use (22) to estimate its covariance matrix. Second, we consider the two-step estimator using the optimal weighting matrix $W_T^{opt} = \hat{S}^{-1}$, where $\hat{S}$ is calculated as (see Newey and West, 1987)

$$\hat{S} = \sum_{j=-k}^{k} \frac{k-|j|}{k}\frac{1}{T}\sum_{s=1}^{T} u_s(\hat{\theta}_1)u'_{s-j}(\hat{\theta}_1), \tag{25}$$

where $\hat{\theta}_1$ denotes the one-step GMM estimator and k is the number of lags. This estimator allows for autocorrelation up to order k in the sample moments and is used because we employ weekly data while the shortest maturity available is equal to one month.

If the moment restrictions imposed by the model are correct, one can expect that the corresponding sample moments, evaluated at the GMM estimate $\hat{\theta}$, are sufficiently close to zero. Within the GMM framework it is easy to test this by means of an overidentifying restrictions test. If the test rejects, one has to conclude that the observed data are inconsistent with the joint validity of all moment conditions. For the one-step estimator $\hat{\theta}_1$, the test statistic is calculated as

$$\xi_1 = g_T(\hat{\theta}_1)'[V\{g_T(\hat{\theta}_1)\}]^+ g_T(\hat{\theta}_1); \qquad (26)$$

where "+" denotes a pseudo-inverse and $V\{g_T(\hat{\theta}_1)\}$ is the (singular) covariance matrix of the sample moments (evaluated at $\hat{\theta}_1$); see Cochrane (1996) for details. In the case of the efficient two-step estimator $\hat{\theta}$, the test statistic can simply be calculated as

$$\xi_2 = T g_T(\hat{\theta})' W_T^{opt} g_T(\hat{\theta}). \qquad (27)$$

Both test statistics have – under the null hypothesis that all moment conditions are valid – an asymptotic Chi-squared distribution with $R - K$ degrees of freedom, where $R - K$ corresponds to the number of overidentifying restrictions (i.e. the number of moments minus the number of unknown parameters). The overidentifying restrictions test serves as a general misspecification test of the model in the sense that the model assumptions are tested against a general unspecified alternative. This is important, because most recent studies using both dimensions of the term structure data, use maximum likelihood to estimate the model parameters. Not only does this require the imposition of much stronger assumptions upon the measurement errors (typically including normality, homoskedasticity and restrictive correlation structures), it also limits the search for misspecifications, if any, to well-specified alternatives.

4 Data and Summary Statistics

Empirical results for term structure models for countries other than the USA are relatively scarce (see Brown and Schaefer, 1986, for an application using UK data). In this paper we use data from four different countries. This way we can compare the results across countries and answer the question whether the relative failure of empirical term structure models for the USA extends to other countries. In order to make a solid comparison between different countries, data that are comparable across countries have to be used. While zero-coupon data are perhaps readily available for the USA, corresponding data sets for other countries are more difficult to obtain.[3]

[3] Moreover, in countries like Germany and Japan, investors have to pay taxes on interest income on government bonds. This implies that government bonds issued in Germany and Japan

To ensure that the data we use are measured as similarly as possible we construct the yield curves as they are implied by LIBOR (London Interbank Offer Rate) and swap rates. Details about the construction of the yield curve data are given in Appendix C. Note that our data sets do not include zero coupon government bond yields, because the interbank LIBOR rates involve a small but nonzero default risk. As described in Hull (2000, p. 125-127) standard practice for relatively small maturities is to set the LIBOR rate equal to the Treasury note rate plus an additional premium of about 30 basis points. This swap spread can vary however, according to demand and supply conditions in the market. Note that modeling the LIBOR yield curve may be more relevant than the zero coupon yield curve for government bonds, since most interest rates derivatives are priced using these LIBOR rates. Several other studies, including Honoré (1997), Duffie and Singleton (1999), Dai and Singleton (1999) and Piazzesi (2000), also use data on LIBOR and swap rates.

The data we use are taken from Datastream, and yield curves are constructed on the basis of maturities of 1, 3, 6, 12, 24, 36, 48 and 60 months. The data are on a weekly basis, covering the time span April 7, 1987 to March 23, 1999 (625 observations) for Germany, the UK and the US. For Japan, swap rates were only available since September 19, 1989, so we employ a smaller sample containing 497 observations.

Table 1 presents some summary statistics of the data. The table shows that the average yield curve for Germany, Japan and the USA is upward sloping, while for the UK the curve is reasonably flat (with a small inverted hump). For the USA, mean yields rise from 6% to 7.3%; and for Germany from 5.8% to 6.5%. The average yields for Japan are much lower, ranging from 3.1% to 4.0%. For each of the countries the volatility, as measured by the sample standard deviation, is decreasing with maturity, which makes the estimated volatility curves downward sloping. Compared to the other countries, the slope of the curve for the USA is relatively small.

In Figure 1 graphical presentations of the four data sets are provided. The figures clearly stress the large time-series variation of the yield curves. Average yield curves are fairly low in the last few years of the sample, but very high in the early 1990s. This is the case for each of the data sets, although the level and the variation in the average yield curve differ across the countries. The figures clearly indicate that the sample averages of the yield curves are highly dependent on the

are not comparable with, for example, US government bonds. The swap markets are unaffected by withholding of taxes and are sometimes more actively traded (see Dai and Singleton, 1999).

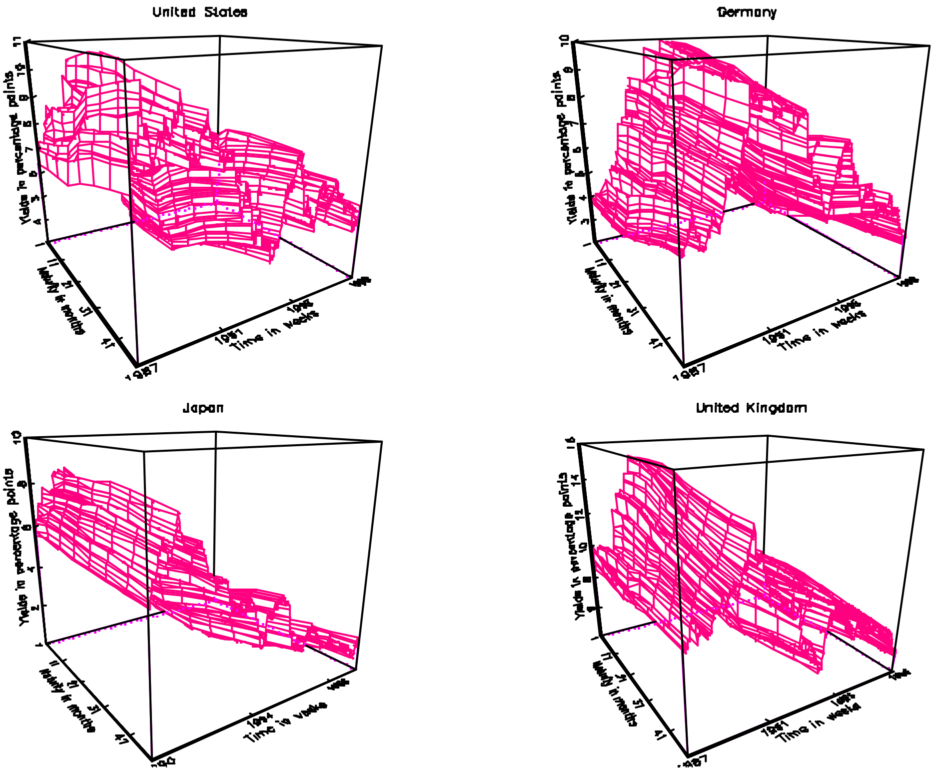

Figure 1: Graphical representation of the yields data

Table 1: Sample means and standard deviations

Maturity (n)	United States	Germany	Japan	United Kingdom
1	0.0599 (0.0180)	0.0576 (0.0231)	0.0314 (0.0276)	0.0887 (0.0315)
3	0.0605 (0.0177)	0.0579 (0.0230)	0.0311 (0.0273)	0.0887 (0.0307)
6	0.0610 (0.0173)	0.0577 (0.0226)	0.0308 (0.0266)	0.0878 (0.0294)
12	0.0621 (0.0166)	0.0572 (0.0213)	0.0307 (0.0255)	0.0863 (0.0268)
24	0.0666 (0.0164)	0.0598 (0.0199)	0.0329 (0.0244)	0.0865 (0.0237)
36	0.0693 (0.0156)	0.0615 (0.0178)	0.0355 (0.0231)	0.0872 (0.0217)
48	0.0713 (0.0151)	0.0632 (0.0160)	0.0379 (0.0218)	0.0875 (0.0202)
60	0.0730 (0.0146)	0.0647 (0.0148)	0.0397 (0.0206)	0.0878 (0.0194)
nobs.	625	625	497	625

Average (annualized) yields for maturities of 1, 3, 6, 12, 24, 36, 48 and 60 months, and sample standard deviations, April 7, 1987 – March 23, 1999 (September 19, 1989 – March 23, 1999 for Japan).

exact sample period that is chosen. By means of the chosen moment conditions (see Appendix B), the GMM estimation procedure used in this paper pays attention to the average yield curves over the full sample period. However, it should be stressed that different sample periods may result in quite different estimates.

5 Empirical Results

The generalized method of moments requires the numerical optimization of a quadratic form in the sample moments, which are highly nonlinear functions of the structural parameters of the model. This makes estimation a complicated numerical exercise. A particular problem we encountered was the separate identification of λ_i and σ_i, $i = 1; 2$: To speed up convergence of the numerical optimization procedure and to obtain good starting values for the one-step GMM estimator, we first estimated a reparametrized version of the model, where we estimated $\alpha_i \equiv \frac{1}{2}\lambda_i^2\sigma_i^2$ and λ_i; rather than σ_i and λ_i separately: Using these estimates we calculated starting values to obtain the estimates of the parameters of interest that are presented below.

Moreover, to ensure that the reported estimates indeed correspond to a global minimum of the objective function we re-started the algorithm for a wide variety of starting values and we drew several one-dimensional graphs of the objective function as a function of one of its arguments. As a result, we are fairly confident that the reported estimates correspond to a global minimum. In estimation, the parameters corresponding to the measurement error variances are restricted to be non-negative. For each of the one-month measurement error variances and for some three-month variances, we obtained estimates at the boundary of zero. In these cases, we fixed these parameter values at zero and re-estimated the other parameters.

5.1 Results for the two-factor CLM model

As stressed by Cochrane (1996), the two-step GMM estimator using the optimal weighting matrix, may involve extreme weights for one or more (linear combinations of) the moment conditions. While this produces asymptotically the most efficient GMM estimator, it may lead to estimators with inferior small sample properties or to estimates that are economically less interesting. That is, the estimation procedure may attach a high weight to a combination of moments that is statistically measured very accurately, but economically not very interesting. Ideally though, we expect that the two-step estimates are reasonably close to their one-step counterparts. Below we shall present results for both the one-step and two-step procedures.

The estimation results for the CLM model, using the one-step GMM estimator[4] are presented in Table 2. As described above, the standard errors are corrected for the presence of heteroskedasticity and autocorrelation up to four lags. Table 2 shows that the estimates for the parameters describing the process for the first factor (μ_1; ϕ_1 and σ_1) are not significant for each of the countries, while most of the estimates for the second factor are significant. Estimates for the second factor's mean reversion parameter ϕ_2, vary between 0.8341 and 0.8607. This indicates that the second factor exhibits a relatively slow mean reversion with a half life[5] of about 4 to 5 months. This is in contrast with the estimates of the first factor's mean reversion parameter ϕ_1; that vary between 0.0078 for Japan (indicating a half life of roughly half a week) and 0.6078 for the USA (a half life of almost one and a half month).

The estimated means for the second factor, μ_2; vary between 0.0299 and 0.0886 and are very close to the sample means of the one-month yields as shown in Table 1. This holds for each of the countries. The unconditional variances of the two

[4] The one-step GMM estimator uses the identity matrix as the weighting matrix and employs the moment conditions as derived in Appendix B.

[5] The half life of a factor is calculated as $\ln(0.5)/\ln|\phi_i|$.

Table 2: One-step GMM estimation results of the CLM model

Parameters	United States		Germany		Japan		United Kingdom	
ϕ_1	0.6078	(5.9254)	0.0502	(165.57)	0.0078	(132.42)	0.5182	(5.5749)
ϕ_2	0.8607*	(0.0246)	0.8341*	(0.0250)	0.8404*	(0.0225)	0.8956*	(0.0277)
μ_1	0.0280	(0.1970)	0.0322	(4.1724)	0.0555	(5.9257)	0.0090	(0.0067)
μ_2	0.0588*	(0.0106)	0.0580*	(0.0235)	0.0299	(0.0258)	0.0886*	(0.0031)
σ_1	0.3593	(4.1487)	0.7644	(116.64)	0.7753	(93.3397)	0.7535	(4.2995)
σ_2	0.0429*	(0.0034)	0.0622*	(0.0057)	0.1014*	(0.0349)	0.0532*	(0.0087)
λ_1	-3.8742	(43.545)	-1.8457	(280.37)	-1.8207	(218.43)	-1.8767	(10.707)
λ_2	0	(fixed)	0	(fixed)	0	(fixed)	0	(fixed)
ω_1	0	(fixed)	0	(fixed)	0	(fixed)	0	(fixed)
ω_3	0.0037	(0.0057)	0.0054	(0.0036)	0	(fixed)	0	(fixed)
ω_6	0.0095	(0.0062)	0.0144*	(0.0037)	0.0154*	(0.0019)	0.0111*	(0.0033)
ω_{12}	0.0115	(0.0118)	0.0164*	(0.0075)	0.0211*	(0.0032)	0.0179*	(0.0029)
ω_{24}	0.0160	(0.0151)	0.0202	(0.0147)	0.0234*	(0.0067)	0.0196	(0.0201)
ω_{36}	0.0134	(0.0139)	0.0165	(0.0163)	0.0222*	(0.0078)	0.0193**	(0.0101)
ω_{48}	0.0136	(0.0085)	0.0149	(0.0114)	0.0211*	(0.0060)	0.0189*	(0.0057)
ω_{60}	0.0152*	(0.0050)	0.0156*	(0.0072)	0.0206*	(0.0042)	0.0191*	(0.0037)
χ^2_{df}	78.449	0.0000	90.108	0.0000	54.578	0.0000	68.407	0.0000
df	6		6		7		7	

Notes: *;** denote significance at the 5%, 10% confidence level, respectively. Standard errors are shown in parentheses, except for the overidentifying restrictions test, where the p-value is given. The column "df" indicates the degrees of freedom of the overidentifying restrictions test. Parameters indicated by "fixed" were fixed at 0. The estimation results were obtained using one-step GMM with the 20 unconditional moment conditions described in the appendix.

factors depend upon σ_i; μ_i and ϕ_i (see Appendix B) and the estimates imply that the first factor is much less volatile than the second factor. All estimates of the market price of risk parameter have the expected negative sign. This implies that the average risk premium in this model is positive. Note however that the estimates are insignificantly different from zero for each of the countries. Furthermore, the estimates for the standard deviations of the measurement errors are fairly large and, in general, increase with maturity, while quite a few of them (especially for Japan and the UK) are significant at the 5% level. Even in the case of very weak distributional assumptions, the measurement errors appear to play an important role in fitting the term structure model to the empirical data. This is a clear indication that the model does not perform very well empirically. Another way to evaluate the estimated models is by means of an overidentifying restrictions

Table 3: Two-step GMM estimation results of the CLM model

Parameters	United States		Germany		Japan		United Kingdom	
ϕ_1	0.5787	(3.9822)	0.0300	(22.137)	-0.0148	(23.631)	0.3696	(18.495)
ϕ_2	0.8323*	(0.0298)	0.8284*	(0.0174)	0.8613*	(0.0206)	0.9007*	(0.0213)
μ_1	0.0299	(0.1623)	0.0471	(0.8642)	0.0899	(1.806)	0.0095	(0.0613)
μ_2	0.0567*	(0.0086)	0.0530*	(0.0056)	0.0200*	(0.0067)	0.0873*	(0.0024)
σ_1	0.3597	(2.7493)	0.7548	(15.579)	0.7674	(16.659)	0.7255	(13.584)
σ_2	0.0442*	(0.0051)	0.0622*	(0.0027)	0.0882*	(0.0125)	0.0499*	(0.0054)
λ_1	-3.8649	(28.643)	-1.8689	(38.384)	-1.8400	(39.820)	-1.9480	(36.437)
λ_2	0	(fixed)	0	(fixed)	0	(fixed)	0	(fixed)
ω_1	0	(fixed)	0	(fixed)	0	(fixed)	0	(fixed)
ω_3	0.0055	(0.0044)	0.0059*	(0.0016)	0.0058*	(0.0015)	0.0080*	(0.0029)
ω_6	0.0102*	(0.0049)	0.0141*	(0.0011)	0.0134*	(0.0013)	0.0136*	(0.0030)
ω_{12}	0.0120**	(0.0071)	0.0160*	(0.0015)	0.0179*	(0.0017)	0.0183*	(0.0022)
ω_{24}	0.0153*	(0.0072)	0.0191*	(0.0020)	0.0199*	(0.0021)	0.0198	(0.0192)
ω_{36}	0.0131*	(0.0061)	0.0159*	(0.0021)	0.0193*	(0.0020)	0.0195**	(0.0105)
ω_{48}	0.0131*	(0.0039)	0.0145*	(0.0016)	0.0187*	(0.0018)	0.0190*	(0.0061)
ω_{60}	0.0143*	(0.0025)	0.0151*	(0.0012)	0.0182*	(0.0016)	0.0188*	(0.0039)
χ^2_{df}	80.990	0.0000	87.344	0.0000	36.876	0.0000	49.760	0.0000
df	6		6		6		6	

Notes: $^{*};^{**}$ denote significance at the 5%, 10% confidence level, respectively. Standard errors are shown in parentheses, except for the overidentifying restrictions test, where the p-value is given. The column "df" indicates the degrees of freedom of the overidentifying restrictions test. Parameters indicated by "fixed" were fixed at 0. The estimation results were obtained using two-step GMM (with the optimal weighting matrix (25)), with the results of Table 2 as starting values for the parameters and the 20 unconditional moment conditions described in the appendix.

test. For each of the four countries, this test indicates a sound rejection of the CLM model, which implies that the CLM model is insufficiently flexible to explain the imposed moment conditions, even if measurement errors are allowed for. We discuss the probable causes for these results in section 6 below.

In contrast with the one-step GMM estimator, the two-step GMM estimator uses an optimal weighting matrix (see Hansen, 1982) and puts a larger weight on those (linear combinations of) moments that are statistically more accurate. Asymptotically this can be shown to lead to more efficient estimators of the model parameters.[6] The two-step estimation results are reported in Table 3. We can

[6] The two step GMM estimator is asymptotically efficient within the class of estimators that exploits the set of moment conditions as described in Appendix B.

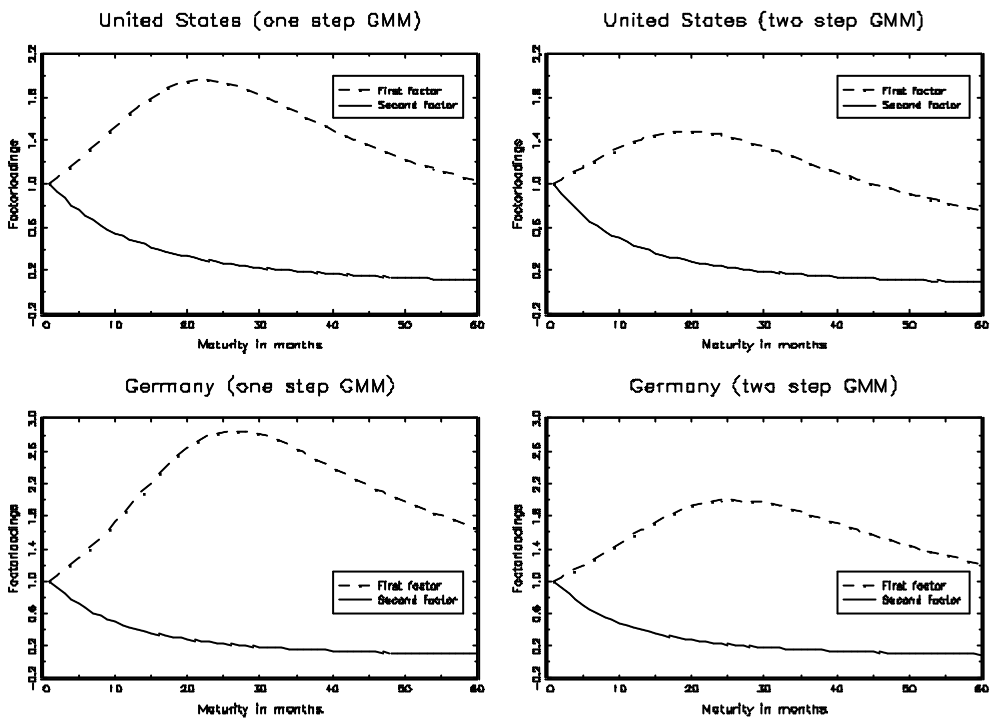

Figure 2: Factor loadings for the CLM model (United States and Germany)

see from the table that the magnitudes of the two-step estimates are more or less comparable to the one-step estimates. In general, the standard errors of the two-step estimates are smaller, which is to be expected. Almost all the estimated standard deviations of the measurement errors are significantly different from zero and, as before, increase with maturity. Considering the overidentifying restrictions tests, we can again clearly reject the CLM model for each of the four countries. Overall, the two-step estimates do not show a more positive picture than the one-step results.

The implied factor loadings b_{1n} and b_{2n} provide information about the statistical behavior of the factors and their impact on the shape of the yield curve. Like in Geyer and Pichler (1998), de Jong (2000), Duarte (2000) and Piazzesi (2000) we draw these factor loadings for all factors and countries (see Figures 2 and 3). For a clear interpretation of the factors, we ensure that the sum of both factors is equal to the one-month yield and therefore we scaled all factor loadings such that both b_{1n} and b_{2n} are equal to one for the one-month yield ($n = 1$).[7] Moreover,

[7] For the U.K., we plot the scaled first factor loading divided by 100.

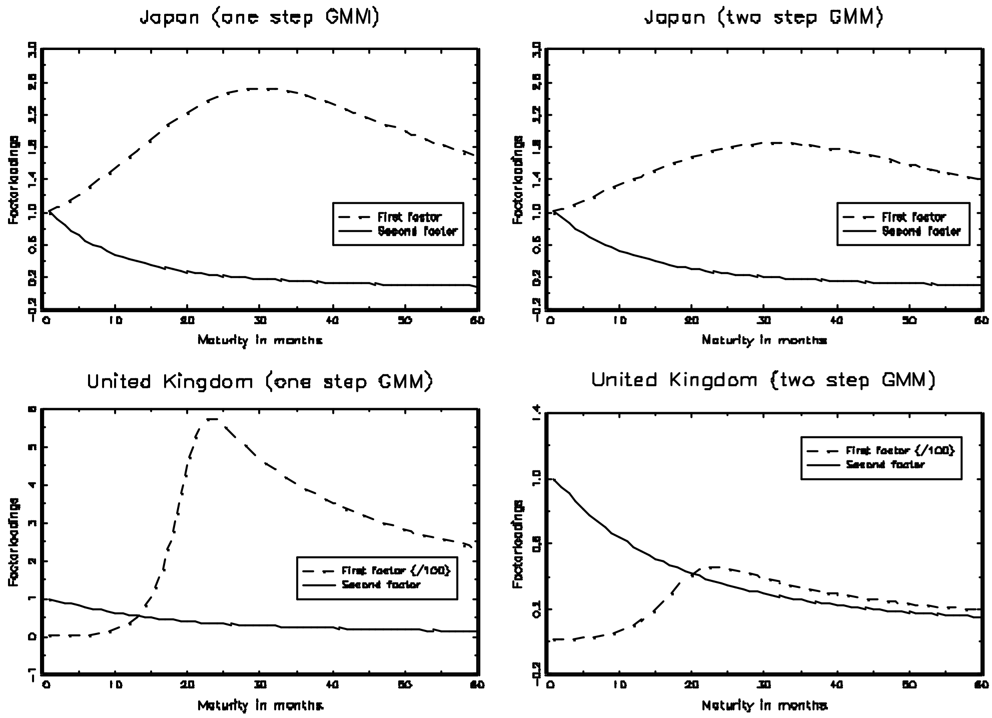

Figure 3: Factor loadings for the CLM model (Japan and United Kingdom)

this scaling eases the comparison of our figures with those reported in the above studies.

From Figures 2 and 3 it follows that the coefficient of the first factor is maximal for maturities around 25–30 months. This finding is consistent across countries, although the magnitude of the factor loadings are somewhat different. The hump shape of the curves for the first factor indicate that this factor is mainly responsible for the "curvature" of the yield curve. For the UK this factor seemingly has a huge impact on the yield curve. However, because the unconditional variance of the first factor is very low, only a very small proportion of the variation in the yield curve can be attributed to movements of the first factor. On the other hand, the second factor loadings are gradually decreasing with maturity, so that we interpret this factor to be mainly responsible for the slope of the yield curve.

As explained above, the CLM model is a special case of the two-factor CIR model, obtained by setting λ_2 to zero. In the next subsection we relax this restriction and present empirical results for the two factor CIR model.

5.2 Results for the two-factor CIR model

The empirical performance of the two-factor CLM model is unsatisfactory for two reasons. First, the estimated standard deviations for the measurement errors are unreasonably large, compared to the sample standard deviations of the yields as shown in Table 1. Second, the overidentifying restrictions tests uniformly indicate that the CLM model should be rejected, even when such large measurement errors are allowed for. One reason why the performance of the two-factor CLM model is disappointing could be that it is incorrect by assuming that the price of risk of the second factor is zero ($\lambda_2 = 0$). Therefore, we shall now relax this restriction and estimate a two-factor CIR model where both factors are priced. The estimation results for the one-step and two-step GMM estimates of the two-factor CIR model are presented in Tables 4 and 5, respectively. First, we discuss the one-step estimates of Table 4.

As in the CLM model, the estimates for the first factor parameters, ϕ_1, μ_1 and σ_1 are insignificant, while the estimates for the second factor parameters are significant. The estimates of the ϕ_1 parameter vary between -0.0729 and 0.7464, while the estimates of ϕ_2 vary between 0.8877 and 0.9848. This indicates that, like in the CLM model, the first factor exhibits relatively fast mean reverting behavior. This is in contrast with the slow mean reverting behavior of the second factor. We also find that the estimates for the second factor's mean parameter μ_2 (which varies between 0.0593 and 0.0888) are very close to the sample mean of the one-month yield. Japan is the only exception of this finding. Again, the combined estimates of ϕ_i, μ_i and σ_i imply that the first factor is less volatile than the second factor for each country. Furthermore, for each country the estimate for λ_1; the market price of risk of the first factor, has the expected negative sign, although it is very inaccurately estimated. The estimates for λ_2; the price of risk of the second factor, are positive and significant only for Japan. While this finding is in contrast with previous studies for the USA (see, for example, Chen and Scott, 1993, Duan and Simonato, 1998, Geyer and Pichler, 1998 and de Jong, 2000), when we calculate the average (log) risk premia for all maturities, these are all strictly positive. For the other countries, some of the average risk premia (especially for the shorter maturities) were negative. We suspect that for the UK this probably can be explained by the fact that we are dealing with an inverted average yield curve.

Again, the estimates for the standard deviations of the measurement errors are large and in general increasing with maturity, which is an indication that also the two-factor CIR model is not performing very well. Moreover, almost all

Table 4: One-step GMM estimation results of the CIR model

Parameters	United States		Germany		Japan		United Kingdom	
ϕ_1	0.7464	(1.5575)	-0.0729	(16.266)	0.2008	(15.744)	0.2679	(38.449)
ϕ_2	0.8953*	(0.0364)	0.8877*	(0.0250)	0.9848*	(0.0364)	0.9143*	(0.0260)
μ_1	0.0317	(0.0624)	0.0880	(1.1960)	0.2002	(3.795)	0.0119	(0.1744)
μ_2	0.0586*	(0.0100)	0.0593*	(0.0077)	0.0866	(0.0856)	0.0888*	(0.0032)
σ_1	0.2392	(1.0912)	0.7921	(11.480)	0.5786	(11.033)	0.8053	(27.815)
σ_2	0.0363*	(1.0077)	0.0491*	(0.0037)	0.0500*	(0.0097)	0.0468*	(0.0071)
λ_1	-5.7023	(24.514)	-1.7803	(25.568)	-2.4353	(48.097)	-1.755	(60.591)
λ_2	5.3586	(7.9079)	4.7882	(3.0603)	23.096*	(7.7838)	0.8454	(1.2330)
ω_1	0	(fixed)	0	(fixed)	0	(fixed)	0	(fixed)
ω_3	0.0041	(0.0076)	0.0078*	(0.0016)	0.0064*	(0.0024)	0	(fixed)
ω_6	0.0085	(0.0083)	0.0140*	(0.0016)	0.0119*	(0.0028)	0.0110*	(0.0040)
ω_{12}	0.0095	(0.0142)	0.0152*	(0.0018)	0.0162*	(0.0034)	0.0168*	(0.0028)
ω_{24}	0.0150	(0.0129)	0.0195*	(0.0014)	0.0205*	(0.0032)	0.0196	(0.0321)
ω_{36}	0.0129	(0.0119)	0.0166*	(0.0012)	0.0207*	(0.0026)	0.0190	(0.0184)
ω_{48}	0.0132**	(0.0077)	0.0149*	(0.0010)	0.0202*	(0.0021)	0.0188**	(0.0104)
ω_{60}	0.0147*	(0.0048)	0.0151*	(0.0009)	0.0197*	(0.0017)	0.0191*	(0.0067)
χ^2_{df}	75.830	0.0000	90.748	0.0000	30.537	0.0000	44.970	0.0000
df	5		5		5		6	

Notes: *;** denote significance at the 5%, 10% confidence level, respectively. Standard errors are shown in parentheses, except for the overidentifying restrictions test, where the p-value is given. The column "df" indicates the degrees of freedom of the overidentifying restrictions test. Parameters indicated by "fixed" were fixed at 0. The estimation results were obtained using one-step GMM with the 20 unconditional moment conditions described in the appendix.

the estimates are significant at the 5% level. Furthermore, even with such large measurement errors, the overidentifying restriction indicate that the two-factor CIR model is rejected for each of the four countries.

The two-step GMM estimates of the CIR model are shown in Table 5 and are in general comparable with the one-step estimates. For the USA, the estimates more or less correspond to the two-factor results of Chen and Scott (1993). Again, we find lower standard errors due to the use of a more two-step efficient estimator. Relative to the first factor, the second factor exhibits a slower mean reversion and is more volatile (as can be calculated from the combined ϕ_i, μ_i and σ_i − estimates). Moreover, the estimates for the second factor's parameters are significant, while their first factor counterparts are not, which is probably due to the very low impact of the first factor on the observed variation in the yield curves. The market price

Table 5: Two-step GMM estimation results of the CIR model

Parameters	United States		Germany		Japan		United Kingdom	
ϕ_1	0.8501*	(0.2278)	-0.1212	(5.3411)	0.1760	(8.2967)	-0.3545	(183.74)
ϕ_2	0.9070*	(0.0686)	0.9140*	(0.0231)	0.9994*	(0.0190)	0.9057*	(0.0075)
μ_1	0.0363**	(0.0195)	0.2019	(0.9373)	0.2300	(2.2318)	0.0142	(1.3080)
μ_2	0.0577*	(0.0153)	0.0599*	(0.0085)	1.4221	(45.207)	0.0865*	(0.0044)
σ_1	0.1418	(0.1594)	0.8044	(3.7592)	0.5978	(5.8281)	1.2198	(130.65)
σ_2	0.0329*	(0.0044)	0.0448*	(0.0033)	0.0447*	(0.0078)	0.0476*	(0.0017)
λ_1	-9.1184	(8.9171)	-1.7544	(8.1649)	-2.3590	(22.872)	-1.1589	(202.67)
λ_2	14.999	(10.142)	12.2813*	(4.7773)	31.380*	(7.4718)	0.0576	(2.2701)
ω_1	0	(fixed)	0	(fixed)	0	(fixed)	0	(fixed)
ω_3	0	(fixed)	0.0055*	(0.0017)	0.0060*	(0.0014)	0.0084*	(0.0004)
ω_6	0.0058*	(0.0012)	0.0112*	(0.0016)	0.0098*	(0.0018)	0.0141*	(0.0039)
ω_{12}	0.0066	(0.0042)	0.0125*	(0.0017)	0.0135*	(0.0021)	0.0181*	(0.0511)
ω_{24}	0.0120*	(0.0045)	0.0165*	(0.0013)	0.0171*	(0.0020)	0.0198	(0.0279)
ω_{36}	0.0107*	(0.0047)	0.0143*	(0.0012)	0.0179*	(0.0017)	0.0193	(0.0146)
ω_{48}	0.0113*	(0.0034)	0.0130*	(0.0011)	0.0179*	(0.0015)	0.0188*	(0.0085)
ω_{60}	0.0128*	(0.0021)	0.0133*	(0.0010)	0.0176*	(0.0013)	0.0187*	(0.0055)
χ^2_{df}	77.046	0.0000	78.654	0.0000	22.464	0.0000	49.822	0.0000
df	6		5		5		5	

Notes: *;** denotes significance at the 5%, 10% confidence level, respectively. Standard errors are shown in parentheses, except for the overidentifying restrictions test, where the p-value is given. The column "df" indicates the degrees of freedom of the overidentifying restrictions test. Parameters indicated by "fixed" were fixed at 0. The estimation results were obtained using two-step GMM (with the optimal weighting matrix (25)), with the results of Table 4 as starting values for the parameters and the 20 unconditional moment conditions described in the appendix.

of risk parameters are negative for the first factor and positive for the second, now being significant for Japan as well as Germany. This implies that when we test the restriction $\lambda_2 = 0$ by means of a simple Wald test, it follows that the two-factor CLM model is rejected in favor of the two-factor CIR model only for these latter two countries. Nevertheless, the overidentifying restrictions tests indicate that the small set of moment conditions implied by the CIR model is not supported by the data. Again the models are very clearly rejected.

As before, we present graphs of the implied factor loadings b_{1n} and b_{2n}, scaled such that both b_{1n} and b_{2n} are equal to one for $n = 1$. Figure 4 shows the factor loadings for the USA and Germany as a function of maturity. For the USA it appears that the first factor is mainly responsible for the general level of the yield

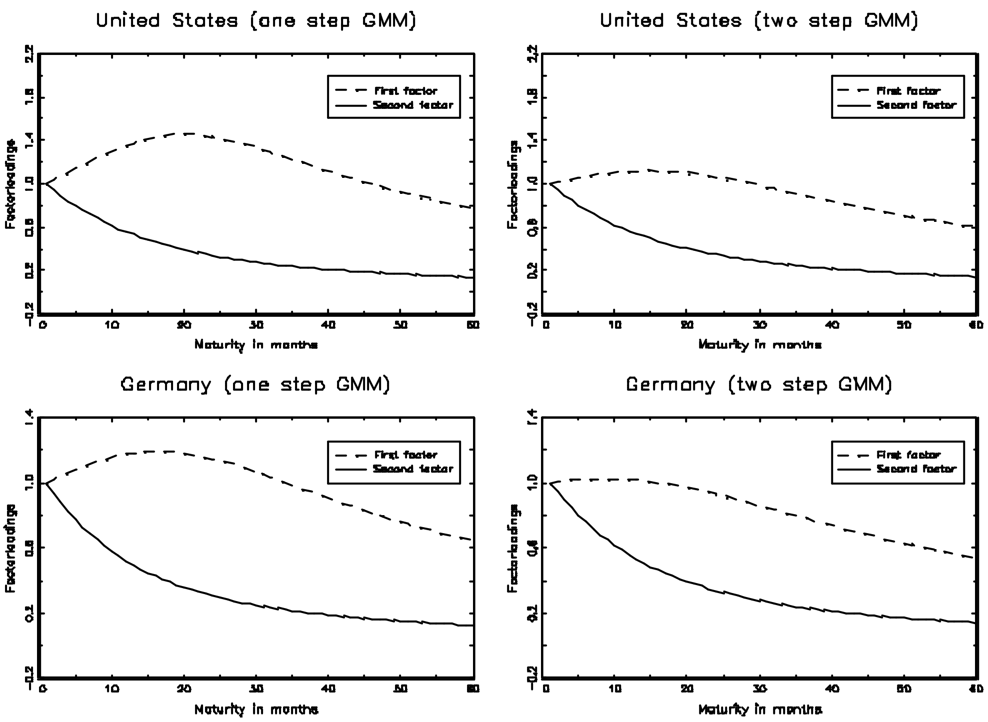

Figure 4: Factor loadings of the CIR model (United States and Germany)

curve, while the second factor is related to the slope of the yield curve. These findings are in line with previous studies that estimate two-factor CIR models for the USA (see Geyer and Pichler, 1998 and de Jong, 2000). For Germany we find both factors to have similar impacts on the yield curve.

Figure 5 presents the factor loadings for Japan and the UK. For Japan the conclusions regarding the two factors are similar to those for the USA and Germany. In contrast, the first factor for the UK appears to determine the curvature of the yield curve, while the second factor is related to the slope of the yield curve. While not as extreme as in the CLM model, the first factor again has a seemingly large impact on the yield curve. When taking into account the low unconditional variance of this first factor, however, the impact on the yield curve is almost negligible.. In general, these findings are in line with the results of the CLM model for the UK.

To summarize our findings, our results show that the two-factor CLM model is unable to satisfactorily explain observed term structure data. The estimated standard deviations for the measurement errors are unreasonably large, while the

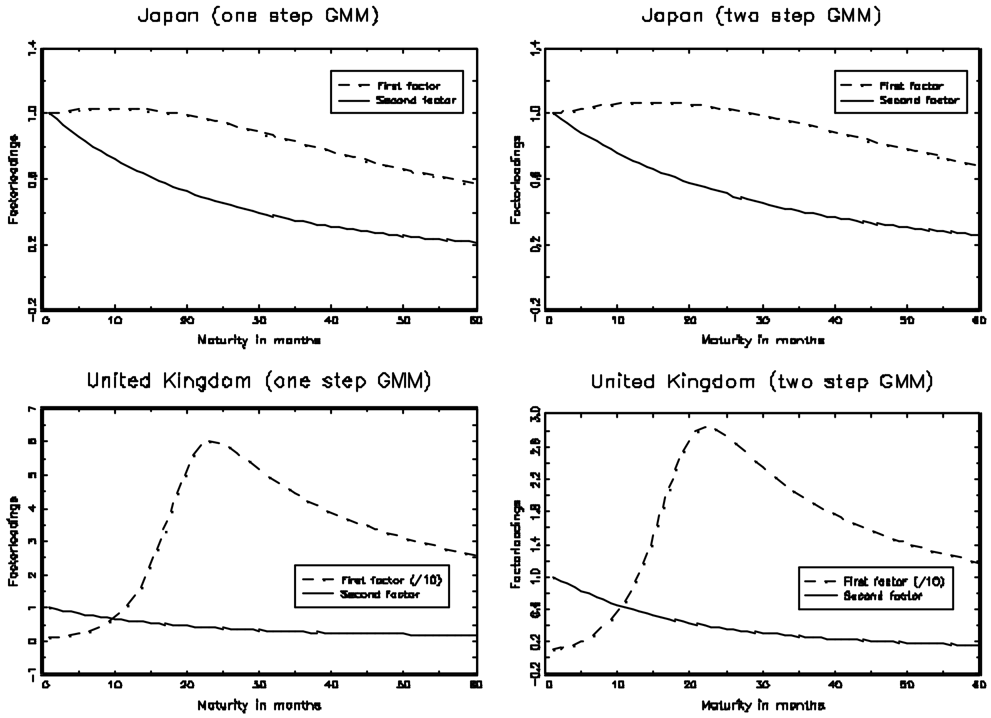

Figure 5: Factor loadings of the CIR model (Japan and United Kingdom)

overidentifying restrictions tests indicate that the model should be rejected. Allowing for two priced factors rather than one and estimating a two-factor CIR model basically does not affect these conclusions. Apparently, the poor performance of these two-factor models as is typically found for the USA (see, for example, Duan and Simonato, 1998, and Geyer and Pichler, 1998) extends to other countries as well (Germany, Japan and the United Kingdom). In the next section we will further explore the question which aspects of the term structure the two-factor models have particular problems with explaining.

6 Explanations for the Statistical Rejections of the Models

The moment conditions that are imposed in estimation (see Appendix B) can be divided into three subsets. A first subset is based upon the unconditional expectations of the yields and corresponds to the average yield curves. A second subset is based upon the unconditional expectations of the squared yields and corresponds,

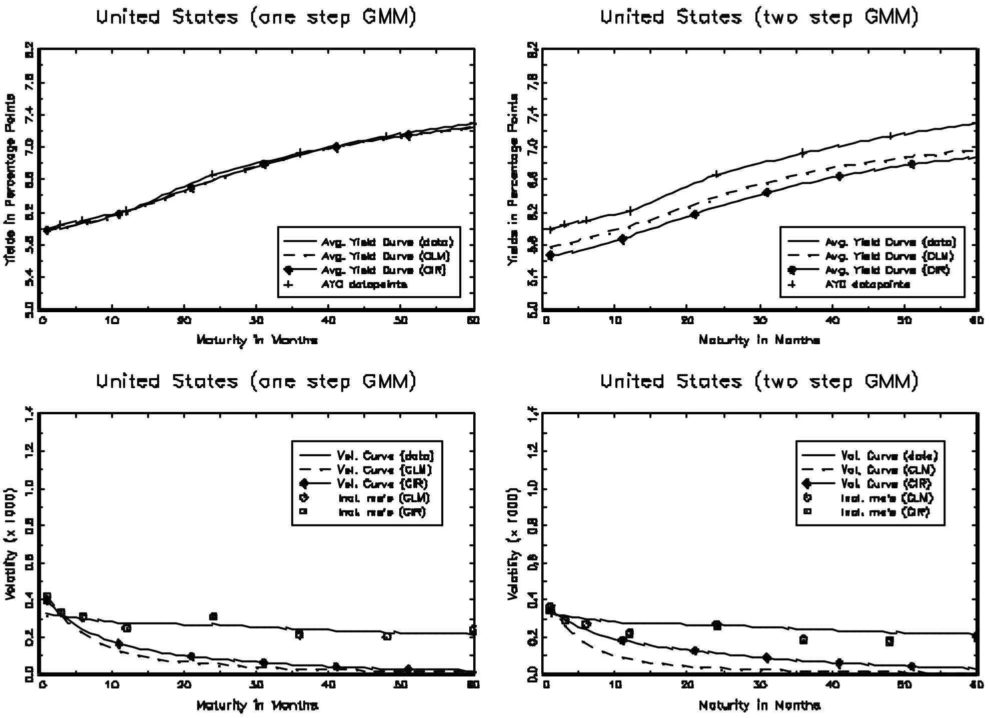

Figure 6: Implied average yield curves and volatility curves for the United States

indirectly, to the volatility curves. Finally, a last subset of moment conditions is related to the dynamic evolution of the two factors and thus to the movement of the yield curves over time. The fact that the overidentifying restrictions tests soundly reject both the two-factor CLM and the two-factor CIR model either indicates that some of the above sample moments are, in an economic sense, quite far from their values implied by the estimated models, or that these sample moments are close to, but nevertheless statistically significantly different from, their estimated counterparts. In this section, we will further explore these alternative explanations and analyze which aspects of the term structure the estimated models have particular problems with.

Let us first consider the upper part of Figure 6. This figure displays the average yield curve for the USA over the period April 1987–March 1999, as well as the unconditional expected yield curves as implied by the parameters estimates of the two-factor models. The figures show that both two-factor models explain the average yield curve reasonably well. For the one-step estimates, the implied yield curves almost exactly match the actual average yield curves, while for the two-step estimates the slope is pretty well captured, though the level of the curve is slightly too high or too low. Previous studies using data for the USA, like Chen and Scott (1993), Geyer and Pichler (1998) and de Jong (2000) find similar results when comparing the implied yield curve with the data. As in Cochrane (1996), the "economic" fit for the one-step estimates is much better than for the two-step estimates.[8]

The lower part of Figure 6 displays the actual and estimated volatility curve. The actual volatility curve is based upon the sample variances of the observed yields, while the estimated curves correspond to the variances of the yields as implied by the estimated two-factor CLM and CIR models, not taking into account the measurement error variances. The implied variances that do take into account the measurement errors are indicated by unconnected points in the figures (circles for the CLM model and squares for the CIR model). Note that the points that include the measurement error variances correspond most closely to the moment conditions that are imposed in estimation. It should be stressed though that the exploited moments are based upon expected squared yields rather than variances of yields. This may have an impact if expected yields are not captured very well by the model.

The lower part of Figure 6 clearly indicates that both two-factor models perform extremely poorly in capturing the actual volatility curves. While the model's volatility curves are downward sloping they substantially underestimate the actual volatilities, except occasionally for the very short end of the yield curve. Also note that the CIR model is able to explain the volatility curve better than the CLM model, which is probably due to the fact that it does not restrict the market price of risk of the second factor to be zero. The figures also explain why the estimates for the standard deviations of the measurement errors, as reported in Tables 2, are so large, at least for the longer maturities.

The estimation results reported in the previous section show that for maturities exceeding three months, the standard deviations of the measurement errors are typically significantly positive and quite large, relative to the standard deviations

[8] The better performance of the one-step estimator in explaining the average yield curve can be understood as follows. The use of the identity matrix in the one-step estimator implies that each moment is equally important. When the two-step estimator is used, with the optimal weighting matrix, it is likely that some (linear combinations of) moments get very high weights in estimation (because these moments are statistically the most accurate ones). While this increases the efficiency of the parameter estimates, as is indeed the case with our results, it may deteriorate the fit of some aspects of the data. In the same spirit, using a GLS estimator in a linear regression model will always provide a lower R^2 than using the OLS estimator. See Cochrane (1996) for additional discussion.

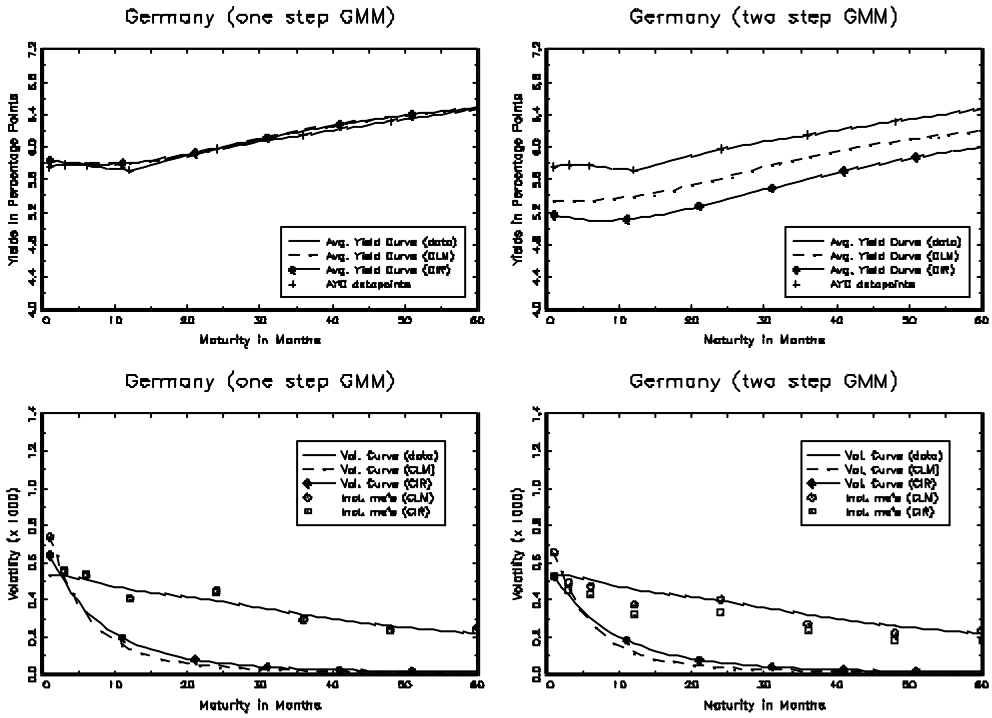

Figure 7: Implied average yield curves and volatility curves for Germany

of the observed yields. Recall that the measurement error variances are restricted to be nonnegative. In some cases boundary solutions were obtained and if this was the case, the standard deviations were set to zero and the other parameters were re-estimated. The …gure explains both results. For the shorter maturities, the estimated volatility curve is very close to its sample equivalent, so the estimated standard deviations of the measurement errors are typically very small or zero. For the larger maturities however, the estimated volatility curve lies considerably below the sample volatility curve. In order to …t the model's moments to the sample moments as imposed by the estimation procedure, the standard deviations of the measurement errors have to be large and this is exactly what we …nd.

The large variances of the measurement errors are an indication that the two-factor models are performing poorly. Apparently, the empirical failure of the two-factor CLM and the two-factor CIR model for the USA cannot be attributed to the typically restrictive assumptions about the distributions of the measurement errors, but is inherent in the insu¢cient ‡exibility of these two-factors models to simultaneously capture both the average yield curve and the volatility curve.

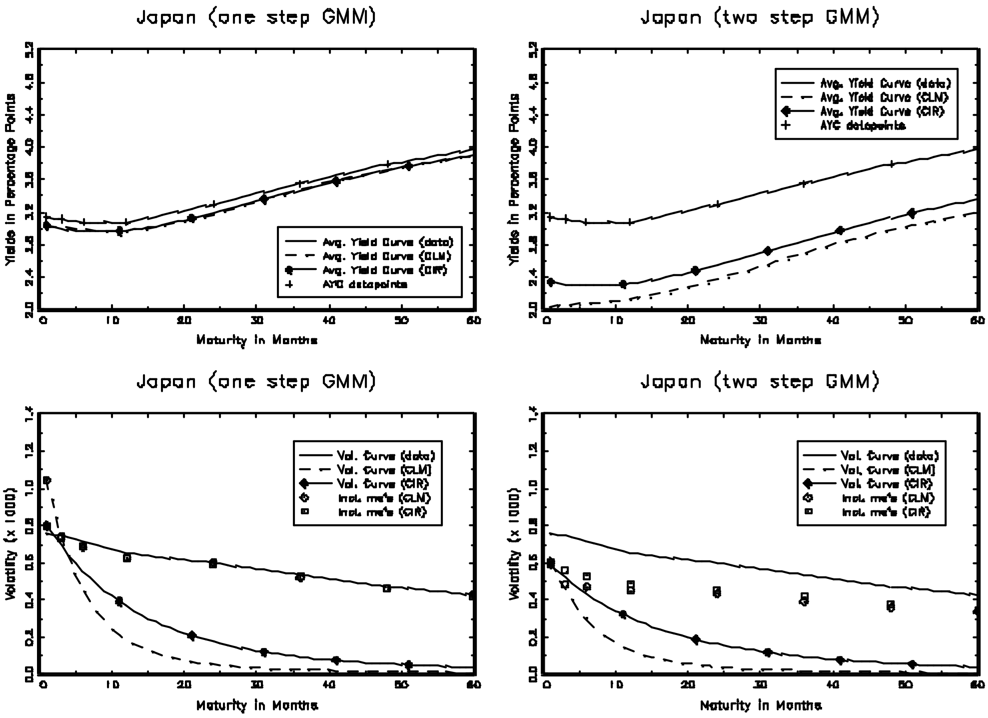

Figure 8: Implied average yield curves and volatility curves for Japan

Figures 7-9 report the actual and estimated yield curves and volatility curves for Germany, Japan and the United Kingdom, respectively. For these countries we also …nd that the one-step estimates do a reasonably good job in approximating the average yield curves, while the two-step estimates provide inferior results, especially for Germany and Japan. For the volatility curves, the situation is no di¤erent than for the USA. Especially for the longer maturities, the volatilities of the yields, as implied by the models, are to a large extent driven by the variances of the measurement errors. These …ndings con…rm the fact that the two-factor a¢ne model does not possess su¢cient ‡exibility to simultaneously capture both the average yield curve and volatility curve of the term structure of interest rates.

Overall, our results indicate that two-factor a¢ne term structure models are very unlikely to provide an accurate description of the yield curve and its development over time. While several studies show that models with three or more factors outperform two-factor models (see, for example, Chen and Scott, 1993, Duan and Simonato, 1998 and Geyer and Pichler, 1998), which is to be expected, there is no reason to believe that these models are overall acceptable. Without exception, the

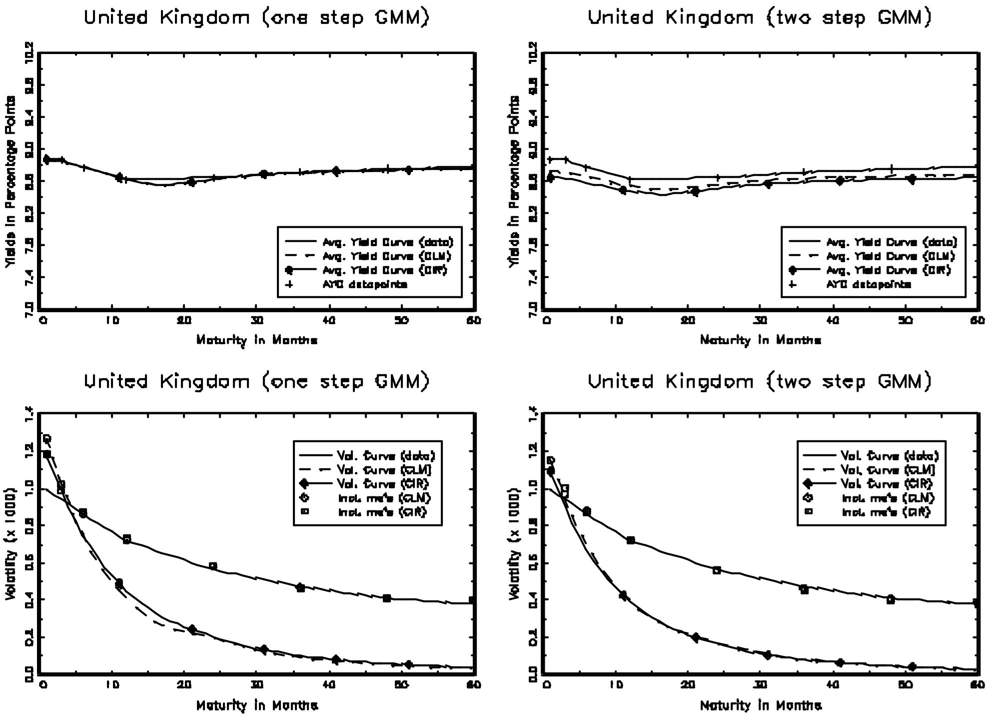

Figure 9: Implied average yield curves and volatility curves for the United Kingdom

extended models are tested against more restrictive versions rather than against more general, unspecified, alternatives. Tests of the latter kind are easily performed in the context of the GMM estimator (with overidentifying restrictions), but are not straightforward in a maximum likelihood context.

7 Conclusions

In this paper we considered the question whether a two-factor affine term structure model is able to describe the term structure of interest rates, while imposing only very weak assumptions on the distributions of the measurement errors on the yields. We estimated two discrete time versions of the Longstaff and Schwarz (1992) two-factor model, for the USA, Germany, Japan and the UK, and found that these models do not possess sufficient flexibility to simultaneously explain multiple aspects of the yield curve. This shows that the poor empirical performance of two-factor affine term structure models for the USA is not due to restrictive and inappropriate assumptions about the measurement errors and, moreover, extends to several other countries. Overall, our results are surprisingly similar across the four countries.

In line with other studies (see, for example, Chen and Scott, 1993, Duan and Simonato, 1998, Geyer and Pichler, 1998, de Jong and Santa-Clara, 1999 and de Jong, 2000), it is found that the model is very well capable of explaining the average yield curve for each of the countries. In contrast, while imposing only very weak assumptions upon the distributions of the measurement errors, we find very large estimates for the standard deviations of these measurement errors which is an indication of poor performance of the model. Indeed, when looking at the volatility curves, we see that measurement error variances are mainly responsible for the shape and location of these curves. Furthermore, the overidentifying restrictions tests show that both the models have to be rejected statistically for each country, even if such large measurement errors are allowed for.

To improve the empirical fit of term structure models, models will have to be constructed that possess more flexibility to explain the volatility structure of the yield curve. While a simple extension of the number of factors, could provide a better explanation of the volatility curve, several recent studies suggest other extensions that may be helpful, see, for example, Duarte (2000), Piazzesi (2000) and Ahn, Dittmar and Gallant (2000).

Appendix A: Derivations of the factor loadings

Consider the price function (11), which implies that the two terms on the right hand side of (7) are

$$
\begin{aligned}
E_t\{m_{t+1} + p_{n-1,t+1}\} &= -x_{1t} - x_{2t} - A_{n-1} - B_{1,n-1}(1-\phi_1)\mu_1 - B_{1,n-1}\phi_1 x_{1t} \\
&\quad - B_{2,n-1}(1-\phi_2)\mu_2 - B_{2,n-1}\phi_2 x_{2t}
\end{aligned} \tag{28}
$$

and

$$
\begin{aligned}
& Var_t\{m_{t+1} + p_{n-1,t+1}\} \\
={}& Var_t\{\sigma_1\sqrt{x_{1t}}(\lambda_1 + B_{1,n-1}) + \sigma_2\sqrt{x_{2t}}(\lambda_2 + B_{2,n-1})\} \\
={}& x_{1t}(\lambda_1 + B_{1,n-1})^2\sigma_1^2 + x_{2t}(\lambda_2 + B_{2,n-1})^2\sigma_2^2.
\end{aligned} \tag{29}
$$

Thus, if we combine (11), (28) and (29), we obtain

$$
\begin{aligned}
p_{nt} &= -A_n - B_{1n}x_{1t} - B_{2n}x_{2t} \\
&= -x_{1t} - x_{2t} - A_{n-1} \\
&\quad - B_{1,n-1}(1-\phi_1)\mu_1 - B_{1,n-1}\phi_1 x_{1t} \\
&\quad - B_{2,n-1}(1-\phi_2)\mu_2 - B_{2,n-1}\phi_2 x_{2t} \\
&\quad + \frac{1}{2}(\lambda_1 + B_{1,n-1})^2\sigma_1^2 x_{1t} + \frac{1}{2}(\lambda_2 + B_{2,n-1})^2\sigma_2^2 x_{2t}.
\end{aligned} \tag{30}
$$

From these conditions the factor loadings as given in the main text are easily derived.

Appendix B: Moment conditions

As derived in the main text (see (18)), the model can be written as

$$y_t^{obs} = A + Bx_t + v_t$$

where v_t is the vector of measurement errors of all available maturities of interest rate data. We assume that

$$\begin{aligned} E\{v_t\} &= 0;\ V\{v_{it}\} = \omega_i^2;\ i = 1, 2, \ldots, m; & (31) \\ \operatorname{cov}\{v_{1t}, v_{1,t-k}\} &= 0;\ k = 1, 2, 3, \ldots & (32) \end{aligned}$$

Using this assumptions, we derive the moment conditions that we use in the GMM estimation procedure.

The expectations of the yields can be determined as

$$\begin{aligned} E\{y_{nt}^{obs}\} &= E\{a_n + b_{1n}x_{1t} + b_{2n}x_{2t} + v_t\} = \\ &= a_n + b_{1n}E\{x_{1t}\} + b_{2n}E\{x_{2t}\} = \\ &= a_n + b_{1n}\mu_1 + b_{2n}\mu_2. \end{aligned}$$

The first set of moment conditions that is used corresponds to

$$\boxed{E\{y_{nt}^{obs} - [a_n + b_{1n}\mu_1 + b_{2n}\mu_2]\} = 0} \qquad (33)$$

for all available maturities n. Note that $E\{y_{nt}^{obs}\} > 0$, so that $a_n + b_{1n}\mu_1 + b_{2n}\mu_2$ should be positive: This is a restriction we impose in the estimation procedure.

Next, the variance of the yields is determined as

$$\begin{aligned} Var\{y_{nt}^{obs}\} &= Var\{a_n + b_{1n}x_{1t} + b_{2n}x_{2t} + v_t\} = \\ &= b_{1n}^2 Var\{x_{1t}\} + b_{2n}^2 Var\{x_{2t}\} + \omega_n^2; \end{aligned}$$

where the unconditional variances of x_{1t} and x_{2t} are given by

$$\begin{aligned} Var\{x_{1t}\} &= E_x\{Var\{x_{1t}\}\} + Var_x\{E\{x_{1t}\}\} \\ &= E_x\{\sigma_1^2 x_{1,t-1}\} + Var_x\{(1-\phi_1)\mu_1 - \phi_1 x_{1,t-1}\} \\ &= \sigma_{11}^2\mu_1 + \phi_1^2 Var\{x_{1,t-1}\} \\ &= \frac{\sigma_1^2\mu_1}{1-\phi_1^2} \end{aligned}$$

and

$$Var\{x_2\} = \frac{\sigma_2^2\mu_2}{1-\phi_2^2};$$

respectively. Furthermore we know that for any stochastic variable z, it holds that

$$E\{z^2\} = Var\{z\} + (E\{z\})^2.$$

Thus we can derive moment conditions based on $E\{(y_{nt}^{obs})^2\}$ as

$$\boxed{E\{(y_{nt}^{obs})^2 - (a_n + b_{1n}\mu_1 + b_{2n}\mu_2)^2 - b_{1n}^2 \frac{\sigma_1^2 \mu_1}{1-\phi_1^2} - b_{2n}^2 \frac{\sigma_2^2 \mu_2}{1-\phi_2^2} - \omega_n^2\} = 0.} \tag{34}$$

Finally, assuming that $Cov\{v_{1t}, v_{1,t-k}\} = 0$ for all $k = 1, 2, 3, \ldots$ we can derive

$$\begin{aligned}
& Cov\{y_{1t}^{obs}, y_{1,t-1}^{obs}\} \\
= \;& Cov\{a_1 + b_{11}x_{1t} + b_{21}x_{2t} + v_{1t}, a_1 + b_{11}x_{1,t-1} + b_{21}x_{2,t-1} + v_{1,t-1}\} \\
= \;& Cov\{b_{11}[(1-\phi_1)\mu_1 - \phi_1 x_{1,t-1}] + b_{21}[(1-\phi_2)\mu_2 - \phi_2 x_{2,t-1}], \\
& b_{11}x_{1,t-1} + b_{21}x_{2,t-1}\} \\
= \;& b_{11}^2 \phi_1 Var\{x_1\} + b_{21}^2 \phi_2 Var\{x_2\}.
\end{aligned}$$

In general, we have

$$Cov\{y_{1t}^{obs}, y_{1,t-k}^{obs}\} = b_{11}^2 \phi_1^k Var\{x_1\} + b_{21}^2 \phi_2^k Var\{x_2\}.$$

The moment conditions based on the autocovariances are then given by

$$\boxed{E\{y_{1t}^{obs} y_{1,t-k}^{obs} - (a_1 + b_{11}\mu_1 + b_{21}\mu_2)^2 - Cov\{y_{1t}^{obs}, y_{1,t-k}^{obs}\}\} = 0.} \tag{35}$$

We exploit these conditions for $k = 1, 2, 3, 4$.

Appendix C: Constructing the LIBOR yield curve

Use LIBOR rates whenever possible

Consider a bond that pays out its \$1 principal at time T, but also makes payments of varying amounts at times $t + \tau_i$ $(i = 1, \ldots, m)$. The amount of payment made at time $t + \tau_i$ is determined by the LIBOR (London Interbank Offer Rate) rate, set at time t:

$$L_{\tau_i t} = \frac{1}{\tau_i}\left(\frac{1}{P_{\tau_i t}} - 1\right) \tag{36}$$

where $\tau_i \in \mathbb{R}^+$ is the maturity of the LIBOR rate (see Baxter and Rennie, 1996, p. 166) and where interest rates are defined in per annum terms. The actual payment made at time $t + \tau_i$ is $\tau_i L_{\tau_i t} = \frac{1}{P_{\tau_i t}} - 1$.

Up to 1 year, LIBOR rates are readily available, so the yield curve for LIBOR rates for $\tau_i \leq 1$ is known for all t (take for example, $\tau_i = i \cdot \alpha$ where α is a fixed interval of one month ($\alpha = \frac{1}{12}$) and take $i = 1, \ldots, 12$). Moreover, given the convention that these assets are denoted in discrete time we have that the price of any zero coupon bond with maturity τ_i is given by:

$$P_{\tau_i t} = \frac{1}{1 + \tau_i L_{\tau_i t}}. \tag{37}$$

From the price of these bonds we can define the observed continuously compounded interest rates for maturity τ_i, $y_t^{obs}(\tau_i)$ as:

$$y_t^{obs}(\tau_i) = -\frac{1}{\tau_i} \ln P_{\tau_i t}. \tag{38}$$

Note that we will need these continuously compounded rates for discounting purposes.

Filtering Bond Prices from Swap rates

The swap rate

A plain vanilla swap is a contract that simply exchanges a stream of varying payments for a stream of fixed amount payments (or vice versa). Typically in such a contract one agrees on receiving a regular sequence of fixed amounts (the swap-rate, $S_{\tau_m t}$) while paying amounts depending on the prevailing LIBOR interest rates, at each payment date.

Suppose the payment dates are $t + \tau_1, t + \tau_2$, ..., $t + \tau_{m-1}$, $t + \tau_m$ and the time between pay-outs, measured in years, is δ, implying the identity: $\tau_i = i \cdot \delta$;

$i = 1, \ldots, m$. The ith payment will be determined by the δ-period LIBOR rate, set at time $t + \tau_{i-1}$. Furthermore, the swap-contract pays out a fixed rate $S_{\tau_m t}$ at each time period. The swap contract is similar to a portfolio that consists of a long position in a fixed coupon bond and a short position in a variable coupon bond. Following Hull (2000, p. 120-130) we know that the value $V_f(t; \tau_m; S_{\tau_m t})$ of the fixed coupon bond is

$$V_f(t; \tau_m; S_{\tau_m t}) = \exp\{-\tau_m y_t^{obs}(\tau_m)\} + \delta S_{\tau_m t} \sum_{i=1}^{m} \exp\{-\tau_i y_t^{obs}(\tau_i)\}; \qquad (39)$$

while the value of the variable coupon bond $V_v(t; \tau_1; L_{\tau_1 t})$ is

$$V_v(t; \tau_1; L_{\tau_1 t}) = (1 + \tau_1 L_{\tau_1 t}) \exp\{-\tau_1 y_t^{obs}(\tau_1)\} \qquad (40)$$

which is equal to 1, if we use definitions (37) and (38). At time t (the start of the contract) the contract must have a value equal to zero:

$$V_f(t; \tau_m; S_{\tau_m t}) - V_v(t; \tau_1; L_{\tau_1 t}) = 0. \qquad (41)$$

Using the initial condition (41) and plugging the definitions (37) and (38) into (39), we calculate the fixed swap rate $S_{\tau_m t}$ to be

$$S_{\tau_m t} = \frac{1 - P_{\tau_m t}}{\delta \sum_{i=1}^{m} P_{\tau_i t}} \qquad (42)$$

(see also Baxter and Rennie, 1996, p. 166 - p. 168).

The price of zero coupon discount bonds can in general not be recovered from swap rates because swaps tend to pay out every six months ($\delta = 6/12$), while we only observe swap rates up to and including 5 years, every year. One way to recover the prices of discount bonds is to linearly interpolate the prices of zero coupon bonds with non-integer maturities (measured in years). To fix notation, denote the integer maturities by the even-numbered indices $\tau_2, \tau_4, \ldots, \tau_m$. By construction, the odd-numbered indices denote the non-integer maturities and define the linear approximation for τ_i $(i = 1, 3, 5, 7, 9)$ as

$$P_{\tau_i t} = \frac{\tau_{i+1} - \tau_i}{\tau_{i+1} - \tau_{i-1}} P_{\tau_{i-1} t} + \left(1 - \frac{\tau_{i+1} - \tau_i}{\tau_{i+1} - \tau_{i-1}}\right) P_{\tau_{i+1} t}; \qquad (43)$$

where we assume that $P_{\tau_1 t}$ and $P_{\tau_2 t}$ are known and correspond to the six month and one-year zero-coupon bond, the price of which can be recovered from the corresponding LIBOR rate directly, using (37).

Filtering the Bond Prices and Discount Rates from Swap Rates

One can recover the bond prices for integer maturities from the swap rate yield curve. As stated above, swap contracts tend to pay out every six month, but only swap rates from integer maturities ("whole" years) are available from the financial markets. Using the linear approximation (43) and equation (42), we can derive bond prices for the same maturities as swap rates. Bond prices for the six month and one-year maturity are available from LIBOR rates and using (43) for $i = 3; 5; 7; 9$ together with (42) and $\Delta = \frac{1}{2}$ we derive that for $j = 4; 6; 8; 10$;

$$S_{\tau_j t} = \frac{\frac{1}{2}\left(1 - P_{\tau_j t}\right)}{\left(P_{\tau_1 t} + \frac{3}{2} P_{\tau_2 t} + P_{\tau_j t}\right) + 2\sum_{i=4}^{j-2} P_{\tau_i t}} : \tag{44}$$

Then using simple algebra, we have that

$$P_{\tau_j t} = \frac{2 - S_{\tau_j t}\left[P_{\tau_1 t} + \frac{3}{2} P_{\tau_2 t} + 2\sum_{i=4}^{j-2} P_{\tau_i t}\right]}{2 + S_{\tau_j t}}; \tag{45}$$

which can be used to calculate the bond prices from the swap rates for the maturities 2,3,4 and 5 years ($j = 4; 6; 8; 10$). Continuously compounded interest rates are then obtained with equation (38).

References

Ahn, D.-H., R.F. Dittmar and A.R. Gallant (2000), Quadratic Term Structure Models: Theory and Evidence, paper to be presented at the 2001 AFA meetings, New Orleans.

Aït-Sahalia, Y. (1996), Testing Continuous-Time Models of the Spot Interest Rate, Review of Financial Studies 9, 385 - 426.

Andersen, T.G. and J. Lund (1997), Estimating Continuous-Time Stochastic Volatility Models in the Short-Term Interest Rate, Journal of Econometrics 77, 434 - 377.

Balduzzi, P., S.R. Das, and S. Foresi (1999), The Central Tendency: A Second Factor in Bond Yields, The Review of Economics and Statistics 80, 62-72.

Bams, D. and P. Schotman (1997), A Panel Data Analysis of A¢ne Term Structure Models, LIFE Working Paper #97-43.

Baxter, M. and A. Rennie (1996), Financial Calculus. An Introduction to Derivative Pricing, Cambridge University Press.

Brown, S. and P. Dybvig (1986), The Empirical Implications of the Cox, Ingersoll and Ross Theory of the Term Structure of Interest Rates, The Journal of Finance 41, 617 - 632.

Brown, S. and S. Schaefer (1994), The Term Structure of Interest Rates and the Cox, Ingersoll and Ross Model, Journal of Financial Economics 35, 3-42.

Brenner, R., R. Harjes and K. Kroner (1996), Another Look at Alternative Models of the Short-Term Interest Rate, Journal of Financial and Quantitative Analysis 31, 85 - 107.

Campbell, J.Y. (2000), Asset Pricing at the Millennium, Paper presented at the annual meeting of the American Finance Association, Boston, 6-8 January, 2000.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1997), The Econometrics of Financial Markets, Princeton University Press.

Chan, K.C., A. Karolyi, F. Longsta¤, and S. Sanders (1992), An Empirical Comparison of Alternative Models of the Short-Term Interest Rates, The Journal of Finance 47, 1209-1227.

Chen, R.R. and L. Scott (1993), Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates, Journal of Fixed Income 3, 14-31.

Cochrane, J.H. (1996), A Cross-Sectional Test of an Investment-Based Asset Pricing Model, Journal of Political Economy 104, 572 - 621.

Cochrane, J.H. (1999), Asset Pricing, Manuscript.

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

Dai, Q. and K. J. Singleton (1999), Speci…cation Analysis of A¢ne Term Structure Models, forthcoming The Journal of Finance.

de Jong, F. (2000), Time Series and Cross-Section Information in A¢ne Term Structure Models, Journal of Business and Economics Statistics 18, 300 - 314.

de Jong, F. and P. Santa-Clara (1999), The Dynamics of the Forward Interest Rate Curve: A Formulation with State Variables, Journal of Financial and Quantitative Analysis 34, 131 - 157.

Duan, J-C. and J.-G. Simonato (1998), Estimating and Testing Exponential-A¢ne Term Structure Models by Kalman Filter, forthcoming Review of Quantitative Finance and Accounting.

Duarte, J. (2000), The Relevance of the Price of Risk in A¢ne Term Structure Models, Working Paper, University of Chicago.

Du¢e, D. (1992), Dynamic Asset Pricing Theory, Princeton University Press, Princeton.

Du¢e, D. and R. Kan (1996), A Yield-Factor Model of Interest Rates, Mathematical Finance 6, 379-406.

Du¢e, D. and K.J. Singleton (1997), An Econometric Model of the Term structure of Interest Rate Swap Yields, The Journal of Finance 53, 1287 - 1321.

Geyer, A. and S. Pichler (1998), A State-Space Approach to Estimate and Test Multifactor Cox, Ingersoll and Ross Models of the Term Structure, Working Paper, University of Vienna.

Gibbons, M. R. and K. Ramaswamy (1993), A Test of the Cox, Ingersoll and Ross Model of the Term Structure, The Review of Financial Studies 6, 619 - 658.

Gong, F.F. and E.M. Remolona (1997), A Three-Factor Econometric Model of the U.S. Term Structure, Sta¤ Report Number 19, Federal Reserve Bank of New York.

Gray, S. (1996), Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process, Journal of Financial Economics 42, 27 - 62.

Hansen, L. (1982), Large Sample Properties of Generalized Method of Moments Estimators, Econometrica 50, 1029-1054.

Honoré, P. (1998), Panel-Data Estimation of Non-Linear Term-Structure Models, Working Paper, Aarhus School of Business.

Hull, J. (2000), Options, futures and other derivatives, fourth edition, Prentice-Hall.

Jegadeesh, N. and G.G. Pennacchi (1996), The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures, Journal of Money, Credit and Banking 28, 428 - 451.

Langetieg, T.C. (1980), A Multivariate Model of the Term Structure, The Journal of Finance 35, 71 - 97.

Litterman, R. and J. Scheinkman (1991), Common Factors A¤ecting Bond Returns, Journal of Fixed Income 1, 54 - 61.

Longsta¤, F.A. and E.S. Schwartz (1992), Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, The 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.

Newey, W.K. and K. West (1987), A Simple Positive De…nite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55, 703 - 708.

Pearson, N.D. and T.S. Sun (1994), Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll and Ross Model, The Journal of Finance 49, 1279-1304.

Piazzesi, M (2000), An Econometric Model of the Yield Curve with Macroeconomic Jump E¤ects, Working Paper, Stanford University.

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