APPLICATION OF THE KALMAN FILTER FOR ESTIMATING CONTINUOUS TIME TERM STRUCTURE MODELS: EVIDENCE FROM THE UK AND GERMANY

Rana Chatterjee *

ABSTRACT

The purpose of this paper is to see how the term structure of interest rates has evolved in the sterling and euro treasury bond markets over the period 1999-2003. German bonds have been used as a proxy for euro-denominated bonds. A state-space representation for the single-factor Cox, Ingersoll and Ross (1985) model is employed to analyse the intertemporal dynamics of the term structure. Closed form solutions for the prices of discount bonds are derived such that they are a function of the unobserved instantaneous spot rate and the model's parameters. Quasi-maximum likelihood estimates of the model parameters are obtained by using the Kalman filter to calculate the likelihood function. Results of the empirical analysis show that while the unobserved instantaneous interest rate exhibits mean reverting behaviour in both the UK and Germany, the mean reversion of the interest rate process has been relatively slower in the UK. The volatility component, which shocks the process at each step in time is also higher in the UK as compared to Germany.

Keywords: CIR model, Kalman filter, term structure

JEL code: G15

---

* Department of Economics, University of Glasgow, Adam Smith Building, Glasgow G12 8RT. Email: 9708248c@student.gla.ac.uk,

1 Introduction

Term structure modelling has two distinct, but related dimensions. The first aspect involves the fitting of a zero-coupon yield curve to a set of cross-sectional bond price observations on any given trading day. The depiction of this relationship between the zero-coupon yields or spot interest rates and their term to maturity is known as the term structure of interest rates. Figure 1 displays the UK Gilt yield curve which shows this relationship for the settlement date 28th January 2004 by applying the term structure estimation model of Nelson and Siegel (1987). [1]

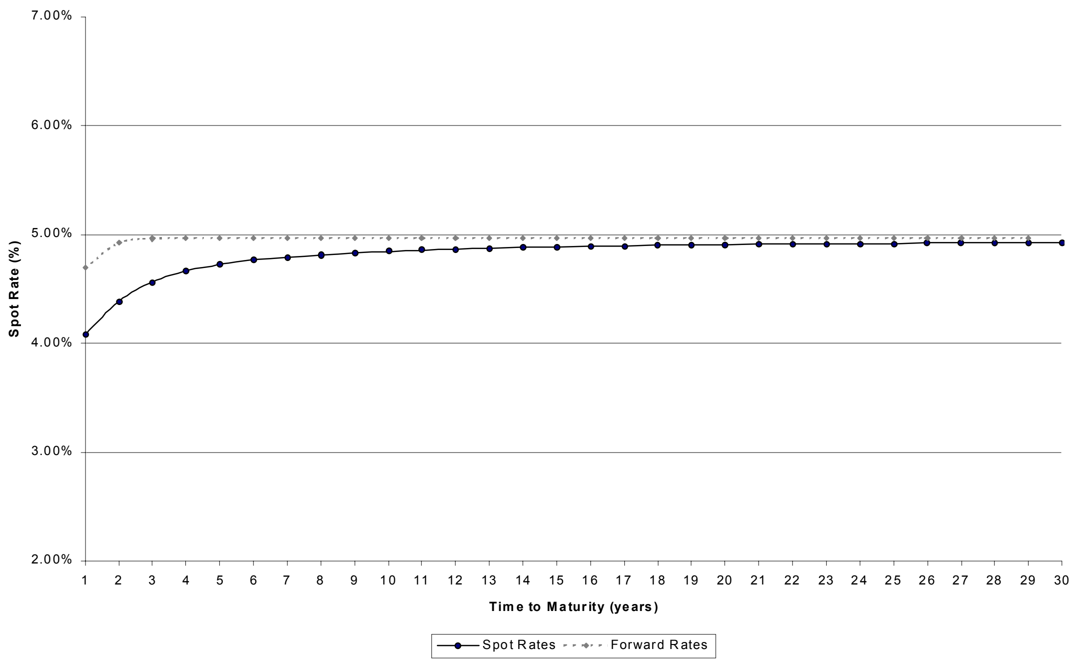

Figure 1. UK Treasury yield curve on January 28, 2004

[1] Nelson and Siegel (1987) model is used to estimate spot interest rates from observable coupon bonds. Market data on bond prices, coupon rates and yield to maturity have been sourced from Datastream.

The second dimension, which is the focus of this paper, relates to the specification of the intertemporal dynamics of the term structure and addresses the issue of how the term structure of interest rates evolves over time. As interest rates are stochastic processes, the models developed in this area rely on the reduction of interest rate uncertainty and attempt to provide parsimonious characterisations of the dynamics of the term structure. Restrictions are imposed on inter-temporal interest rate behaviour by using the no-arbitrage argument. The absence of arbitrage, would ensure that movements of the term structure do not permit conditions to occur under which market participants may earn risk-free profits. There exist various specifications that differ with respect to the number of underlying state variables and the type of the stochastic process. Examples are the models proposed by Vasicek (1997), Cox-Ingersoll-Ross (1985) and Longstaff and Schwartz (1992). However, most modelling approaches are based on the concept that although interest rates change randomly over time, it is possible to divide the change in its value into two parts. The first part is a non-random, deterministic component, called the drift of the process, and the second is the random or noise part which entails the volatility component of the process.

The Vasicek (1977) model is a one-factor partial equilibrium model and starts out with the specification of a time series process for the instantaneous spot interest rate which is treated as the only factor of uncertainty. The no-arbitrage restriction then permits the derivation of a bond pricing formula whereby the bond price is a function of the unobserved instantaneous spot rate and the model's parameters. The approach was extended to include a second factor of uncertainty. Besides the real rate of interest, Richard (1978) chose the expected inflation as the second source of uncertainty. Brennan and Schwartz (1979) model the long rate as the second factor and assume that the short rate is mean reverting to the long rate.

Cox, Ingersoll and Ross (CIR) (1985, CIR hereafter) develop a general equilibrium asset pricing model that allows the deduction of the term structure of interest rates. The model is set up as a single-good, continuous time economy with a single state variable. Multivariate versions are developed by Longstaff and Schwartz (1992) and Chen and Scott (1992). In Longstaff and Schwartz (1992) the two-factors are the short-term interest rate and the variance of changes in the short-term interest rate. Duffie and Kan (1996) define a general class of multifactor affine models of the term structure that allows for the nesting of some of the aforementioned term structure models such as Vasicek (1978) CIR (1985) and Longstaff and Schwartz (1992).

The literature would suggest that three state variables are adequate to explain most of the variability in bond yields. For example, Litterman and Scheinkman (1991) show that this can be captured by the level, the steepness and the curvature of the term structure. This paper focuses on the one-factor CIR model as our empirical estimation showed that the inclusion of additional factors did not increase the performance of the model for either country. A plausible explanation for this could be the limited period of observation. Most studies have concluded that the level is the most important factor in explaining interest variation over time. In fact, Litterman and Scheinkman (1991) have demonstrated that three factors notwithstanding, almost 90 percent of the variation in US Treasury rates is attributable to the variation in the first factor, which is considered to correspond to the level of interest rates. So from an empirical point of view a one-factor CIR model can be considered acceptable.

The purpose of this paper is to see how the term structure has evolved in the sterling and euro treasury bond markets. Euro-denominated bonds are not being placed by the currency area as a whole, but rather by individual countries. A significant feature of the euro capital market is, therefore, the absence of Federal European government debt whose yields would form the natural constituents of euro term structure relationship. In this paper, German bunds have been used as a proxy for euro-denominated bonds as they are seen by market participants as the main component of the euro yield curve. Although there exists a considerable literature on empirically estimating the CIR model, most of the tests have been performed on US data. To the best of my knowledge, this is the first study that estimates this model for the UK and German bond data since the launch of the single currency. By bringing together the empirical findings for the euro and sterling treasury bond markets I attempt to compare the dynamics of their respective term structures. This investigation into the intertemporal behaviour of the euro and sterling term structure could provide evidence on whether there exists any common factors.

The rest of the chapter is organised as follows. Section II provides the theoretical framework that discusses in detail the one-factor CIR model for the instantaneous interest rate. Section III provides an overview of the different estimation methods. In Section IV the state space representation of the CIR model is formulated and, in Section V the Kalman filter algorithm is employed. Section VI presents the data and results. Finally, Section VII concludes.

II Theoretical Framework

Bond pricing in continuous time

Before proceeding further, we will reiterate some of the key bond pricing relationships in a continuous-time setting. We define a pure discount bond as a contract that pays one unit of currency at its maturity date and we denote its value by the function $P(t,T)$. The first argument, *t*, refers to the current time, while the second argument, *T*, represents the pure discount bond's maturity date. It follows that $t < T$. Given the contractual nature of the pure discount bond, $P(T,T) = 1$. In other words, the pure discount bond has a value of £1 at maturity.

Given the pure discount bond price for any given maturity, the associated spot rate of interest for that date can be determined. The spot rate, which we denote as $z(t,T)$, is the continuously compounded rate of return that generates the observed prices of the discount bond. The spot rate can then be solved for as follows:

$$P(t,T)e^{(T-t)z(t,T)} = 1, \tag{1}$$

$$\ln(P(t,T)e^{(T-t)z(t,T)}) = 0,$$

$$\ln(e^{(T-t)z(t,T)}) = -\ln P(t,T)$$

$$z(t,T) = -\frac{\ln P(t,T)}{T-t} \tag{2}$$

As mentioned above, equilibrium models derive a process for the short rate, *r*, and then explore what the process for *r* implies about bond prices. The short rate, *r*, at time t is the rate that applies to an infinitesimally short period of time at time t. It is also referred to as the instantaneous short rate. Bond prices and other derivative prices depend on the process followed by r in a risk-neutral world. The instantaneous short rate cannot be directly observed and is a theoretical construct designed to facilitate the modelling process. However, the way it is determined in a term structure model is significant. For instance, r can be a state variable itself or it can be an affine sum of state variables in the affine framework.

As interest rate processes are stochastic processes, developing an affine term structure model involves a specification of a stochastic process for the state variables, or factors, that drive the dynamics of the term structure. In a one-factor term structure model, the factor is generally taken to be the instantaneous spot rate of interest. It is possible to divide the change in its value into two parts, the first is a non-random deterministic component, called the drift of the process, and the second is a diffusion term or random part, which is the variance of the process. This involves the assumption that the interest rate process is a Markov process[2] and that its dynamics can be described by the following first-order stochastic differential equation:

$$dr = m(r)dt + s(r)dw$$

where w is a standard scalar Wiener process. [3]

---

[2] A Markov process is a particular type of stochastic process where only the present value of a variable is relevant for predicting the future. When Markov processes are considered, the variance of changes in successive time periods are additive.

[3] A Wiener process is a particular type of Markov process with a mean change of zero and a variance rate of 1.0 per year. The drift rate of zero means that the expected value of w at any future time is equal to its current value. The variance rate of 1.0 means that the variance of the change in w in a time interval of length T equals T.

The interpretation of this stochastic differential equation is that the differential change in the short rate is composed of a drift term , which is non-random, and a diffusion term, which is random and includes a differential increment of a Wiener process.

The Cox, Ingersoll, and Ross Model

In the CIR model, a representative agent with constant relative risk aversion faces production opportunities which evolve according to movements in a single state variable, which is in turn described by a first-order differential equation. This implies that the instantaneous interest rate is proportional to the state variable (an thus can be thought of as the state variable itself) and its process is given by:

$$dr = k(\theta - r)dt + \sigma\sqrt{r}dw^P \qquad (3)$$

By virtue of the square root process interest rates are prevented from becoming negative. It also implies that the volatility of the short-term interest increases with an increase in the level of short-term interest rates. dw^P is a Wiener process under the P-measure which is explained below.

With this process for the short rate and the assumptions made concerning preferences, CIR show that the time-*t* yield to maturity on a pure discount bond paying one unit of the consumption good in (*T* - *t*) periods can be written as:

$$R(t,T) = \frac{B(t,T) - \log A(t,T)}{T - t}$$

where

$$B(t,T) = \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + k)(e^{\gamma(T-t)} - 1) + 2\gamma}$$

and

$$A(t,T) = [\frac{2\gamma e^{(k+\gamma)(T-t)/2}}{(\gamma + k)(e^{\gamma(T-t)} - 1) + 2\gamma}]^{2k\theta/\sigma^2}$$

where $\gamma = \sqrt{k^2 + 2\sigma^2}$.

Thus the yield is an affine function of the short rate, and depends upon the long-run level of the short rate θ, the degree of mean reversion κ, the volatility of the short rate σ.

Probabilistic Framework

The prices of bonds are critically dependent upon the probabilities of events occurring in the markets' attitude to risk. This uncertainty in the market is characterized by the probability space (Ω, $\Im$, P) where Ω is the sample space, $\Im$ is the set of all events, and P is the probability measure defined on (Ω, $\Im$). Ω, the state space can be thought of as the set of all possible histories of the world up to some maximum time horizon. An "event" is a subset of Ω which can be given a probability of occurring. $\omega \in \Omega$ could be called a sample point. For example, a contingent claim [4] could be defined as a financial claim that pays its cash flow, or cash-flows, only if some predetermined state of the world ($\omega \in \Omega$) is achieved at some point, or points, in the future. Interest-rate contingent claims include caps, floors, swaptions, and callable bonds.

Absence of arbitrage and equivalent martingale measures

The no-arbitrage condition and martingale measures are concepts of fundamental importance in the analysis of contingent claims. A martingale is a zero-drift stochastic process. [5] A variable, *p,* which we define as the relative price of one security with respect to another, follows a martingale if its process has the form:

$$dp = \sigma dw$$

A martingale has the convenient property that its expected value at any future time is equal to its value today. The *equivalent martingale measure* result shows that, when there are no arbitrage opportunities there exists a pricing kernel (also known as state price deflators or stochastic discount factors) such that the product of the security price and this pricing kernel constitutes a martingale. Therefore, we define a pricing kernel such that the product of the security price and the pricing kernel follows a martingale process.

$$M(t)P(t,T) \equiv E_t^P[M(t)P(t)]$$

$$P(t,T) = E_t^P\left[\frac{M(T)}{M(t)}P(T)\right]$$

[4] A claim whose value depends on the value of another asset.

where $M(t)$ is the state price deflator at time *t*, given exogenously. If we consider zero coupon bonds that yield 1 unit of account at time T this equivalence can be simplified to

$$P(t,T) = E_t^P[\frac{M(T)}{M(t)}]$$

The no-arbitrage assumption poses restrictions on the drift and diffusion terms of the pricing kernel's stochastic process. If we assume that the bond dynamics are described by the following first-order stochastic differential equation:

$$dP(t) = \mu_P(t)dt + \sigma_p(t)dW$$

Then analogously we can write the pricing kernel dynamics as:

$$dM(t) = \mu_M(t)dt + \sigma_M(t)dW(t)$$

By definition, it follows that:

$$E_t^P[dM(t)P(t)] = 0$$

If we ignore the time dependency for convenience we have:

$$E_t^P[Pd(M) + Md(P) + (dM)(dP)] = 0$$

$$P\mu_M + M\mu_P + \sigma_M\sigma_P = 0$$

$$\frac{\mu_M}{M} = -\frac{\mu_P}{P} - \frac{\sigma_M}{M}\frac{\sigma_P}{P}$$

---

[5] Hull (2001)

To assume that the bond is a risk-free bond would be to assume that $\sigma_P = 0$. Moreover, it should earn the risk-free rate of return or the short rate, which would mean that $\mu_P = rP$. This implies that

$$\frac{\mu_M}{M} = -r$$

Using this relationship, we can show that

$$\frac{\sigma_M}{M} = -\frac{\mu_P / P - r}{\sigma_P / P}$$

The absence of arbitrage would, intuitively, mean that assets which exhibit the same risk should earn exactly the same (excess) return. Therefore, the return/risk ratio should be the same for all assets as given by, say, λ. So, by definition

$$\lambda \equiv \frac{\mu_P / P - r}{\sigma_P / P}$$

therefore

$$\frac{\sigma_M}{M} = -\lambda$$

This exercise demonstrates that the no-arbitrage condition restricts the drift coefficient of the pricing kernel dynamics to be the negative of the short rate, and the diffusion coefficient to be the negative of the market price of risk. This can be stated formally as:

$$\frac{dM(t)}{M(t)} = -r(t) - \lambda(t)dW(t)$$

Risk-adjusted Processes

Using these risk-adjusted processes, the effect of the market price of risk corresponding to the short rate on the level of the level of the short can be incorporated in the model. Therefore, if the zero coupon bond prices follow a martingale process the CIR process given by equation (3) can be represented as:

$$dr = (k(\theta - r) - \lambda r)dt + \sigma\sqrt{r}dw^Q \qquad (4)$$

where dw_t^Q is a Wiener process under the Q-measure. λ is the market value of risk - the covariance between changes in the interest rate and the market portfolio.

For the one-factor CIR model, the solution for the nominal price of a pure discount bond is given by

$$P(t,T) = A(t,T)e^{-B(t,T)r} \qquad (5)$$

where, after incorporating the market value of risk, λ,

$$A(t,T) = [\frac{2\gamma e^{(k+\gamma+\lambda)(T-t)/2}}{(\gamma+k+\lambda)(e^{\gamma(T-t)}-1)+2\gamma}]^{2k\theta/\sigma^2} \quad (6)$$

$$B(t,T) = \frac{2(e^{\gamma(T-t)}-1)}{(\gamma+k+\lambda)(e^{\gamma(T-t)}-1)+2\gamma} \quad (7)$$

$$\gamma = \sqrt{(k+\lambda)^2 + 2\sigma^2} \quad (8)$$

The continuously compounded yield for discount bonds is given by:

$$R(t,T) = -\frac{\log P(t,T)}{T-t} \quad (9)$$

Using equation (5), this can be rewritten as:

$$R(t,T) = \frac{-\log A(t,T) + B(t,T)}{T-t} \quad (10)$$

III Estimating the CIR model

A variety of methods have been developed in the finance literature for the estimation of CIR-type models. The two basic approaches may be categorised as the cross section approach and the time series approach.

In the cross-section approach, only information on the yields of bonds at different maturities at a point in time is used in the estimation process. One thereby obtains a different set of parameters for each time period. The state variable r_t, treated as an additional unknown parameter, is estimated jointly with the structural parameters. A shortcoming of this approach is that the parameters are not constant over time and to that extent it contradicts a basic assumption of the underlying economic model. An example of this approach can be found in Brown and Schaefer (1994).

The time series approach, on the other hand, focuses on the dynamic implications of the model and ignores the cross-sectional information. This approach is based on fitting equation (3) to estimate the parameters, using observable data (e.g. the yield of one-month Treasury bills or money market rates) as an approximation of the unknown parameter estimates. An inadequacy of this approach is that it is not possible to obtain an estimate of the market price of risk, λ, which is necessary for valuation purposes. Examples of this approach include Chan, Karolyi, Longstaff and Sanders (1992) and Longstaff and Schwartz (1993).

These shortcomings saw the emergence of a third approach using panel data which takes the dynamic and cross-sectional information into account simultaneously. Gibbons and Ramaswamy (1993) highlighted the need for an estimation process which is based on observed market prices and jointly controls for the assumed factor dynamics. Pearson and Sun (1994) attempted to accomplish this by formulating a likelihood function for a two-factor CIR model on the basis of the conditional density of the underlying factors. The model is estimated by replacing the two factors by two zero-coupon yields that are observed without error. However, the method does not account for the complete cross sectional information since it employs only two points of the term structure. Moreover, the assumption that the two zero-coupon yields are observed without measurement error is open to question.

The affine model posits an exact relation between factors and yields. When using more maturities than factors, this relation cannot be satisfied by all elements of the yield vector as it would give rise to the singularity problem in matrices. Hence, some structure of measurement errors is necessary. In addition to the mathematical requirement for having measurement errors is the economic justification for their inclusion. It is important to note that affine models assume frictionless markets which is obviously not the case in actual practice. Bid-ask spreads, rounding of prices, differences in the timing of observing financial variables, and non-synchronous trading are actual market features that detract from the notion of a frictionless market.

In view of the need for some structure of measurement errors, it is important to make assumptions about this structure. Several alternatives have been put forward in the literature. 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 directly to obtain the factors. The other yields are assumed to be measured with a normally distributed measurement error. It seems more reasonable however to assume that all yields are hit by some measurement errors and this assumption is made by Geyer and Pichler (1999). They further assume that these measurement errors have zero means, an unknown variance and are serially and cross-sectionally uncorrelated. In effect, they assume a diagonal covariance matrix for the measurement errors.

The essence of this discussion is that although yields are observed, these observations are by assumption imperfect due to market frictions of all kinds. The factors supposedly present are unobservable and are to be predicted according to their assumed distribution and the noisy indirect signals that emanate from the resulting yields. Regardless of the structure of the measurement error chosen, given that we cannot simply invert the model to find the unknown factors, a more sophisticated filtering method is needed. The objective is to find a method for filtering out the desired true signal and the unobserved components from this unwanted noise. This where the Kalman filter comes in. It is a recursive algorithm that begins with an educated guess as to the initial values for the state variables. Parameter values can then be estimated using maximum likelihood methods. In the context of bond markets it identifies the unobserved state variables that govern bond price dynamics.

The Kalman filter has been used in a series of papers dealing with the estimation of exponential affine term structure models. Affine term-structure models are constructed by assuming that the bond price is a linear function of the underlying state variables that provide uncertainty in the model. The Kalman filter is a linear estimation method and makes use of this affine relationship between bond prices and state variables to subsequently estimate the parameter set. The main advantage of this technique stems from the fact that it allows the state variables to be unobserved magnitudes.

The nature of the application of the Kalman filter depends on whether the term structure model is Gaussian such as the Vasicek model or non-Gaussian such as the

CIR model. A Gaussian distribution is fully characterised by its first two moments and the exact likelihood function is obtained as a by-product of the Kalman filter algorithm. In the Gaussian case examples are due to Lund (1997) and Babbs and Nowman (1999), who estimate a two-factor generalised Vasicek model. Both Lund and Babbs and Nowman observe eight spot rates with maturities between one and 10 years. When using non-Gaussian models, however, the exact likelihood function is not available in closed-form, but a quasi-maximum likelihood estimator can be constructed from the first and second conditional moments of the state variables. For the non-Gaussian CIR model, examples are due to Duan and Simonato (1995) and again Lund (1997). Duan and Simonato (1995) investigate the general affine case. As an illustration they fit one and two-factor CIR models to rates of maturities up to nine months.

In this paper, a panel-data estimation of the CIR model is presented which draws on the work of Geyer and Pichler (1998). The approach is based on a state-space representation of the term structure model where the underlying state variable(s) is treated as unobservable. This obviates the need to employ proxies for the unobserved factors. The yields are affine in the underlying state variables and the model explicitly allows for measurement errors. Quasi-maximum likelihood estimates of the model parameters are obtained by using an approximate Kalman filter to calculate the likelihood function.

IV The state space representation

In this section we demonstrate the reformulation of the model given by equation (4) in the state space form as explained by Harvey (1989). In this formulation, we define a *measurement equation* that relates the observable, or measurable bond yields to the unobservable state variables. The unobservable state variables, are in turn, assumed to follow a Markov process described by the *transition equation*. Yields on zero-coupon bonds are the inputs to the estimation process. In principle, we require only one zero-coupon rate for each factor used in the estimation. So a one-factor model would require only one zero-coupon bond yield. However, we have chosen eight maturities that span the yield curve from 2 years to 25 years in order to incorporate information affecting trading at the short, medium and long ends of the yield curve. This cross-sectional information is particularly useful in specifying the market price of risk parameter.

In the CIR model, the measurement equation represents the affine relationship between zero coupon bond yields and the state variables. Under the assumption that measurement errors are additive and normally distributed, we have

$$R_t = d(\psi) + Z(\psi)X_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, H(\psi)) \tag{11}$$

where ψ contains the unknown parameters of the model including the parameters from the distribution of measurement errors. Therefore, $\psi = (\theta, k, \sigma, \lambda, h_{1,\cdots N})$. The error terms ε_t are measurement errors to allow for noise in the sampling process of the data. The variance-covariance matrix of the measurement errors is assumed to have the form $H = h_1, \cdots h_N$ along the diagonal. In this estimation, 8 different maturities are being considered. Therefore, the variance-covariance matrix of the measurement errors, *H*, is an 8 x 8 diagonal matrix.

$$H = \begin{bmatrix} h_1^2 0 \cdots 0 \\ 0 h_2^2 \cdots 0 \\ \vdots \vdots \ddots \vdots \\ 0 0 \ldots h_8^2 \end{bmatrix}$$

The values in the diagonal would differ implying that the variance of measurement errors will depend on the maturities under consideration. This can be justified on the grounds that trading activity and, therefore, bid-ask spreads are not equally distributed across maturities. Therefore, the assumption is made that measurement errors in bond yields are additive and normally distributed. In the case of a one-factor model, equation (11) would read as:

$$R_t = A_0 + A_1 X_t + \varepsilon_t \qquad \varepsilon_t \sim N(0, H(\psi))$$

where $A_0 = -\dfrac{\log A(t,T)}{T-t}$, $A_1 = \dfrac{B(t,T)}{T-t}$.

The stochastic differential equation (4) represents the dynamics of the state variable as specified in continuous time. As the transition equation captures the discrete dynamics of the state variable, it corresponds to the discrete time version of equation (4). This along with a first order autoregression model are used to formulate the transition equation,

$$X_t = d(\psi) + \phi(\psi) X_{t-1} + u_t, \qquad u_t / \Im_{t-1} \sim N(0, Q_t,) \tag{12}$$

where

$d(\psi) = \theta_j (1 - e^{-k\Delta t})$ and $\phi = e^{-k\Delta t}$

$\Delta t =$ the time interval in the discrete sample (here 1 week)

and so the discretisation step $\Delta t = 1/52$ for weekly data.

It is important to note that the matrix Q_t is diagonal and is dependent on the state of the process. For a three-factor model, the conditional variance of the transition system would have the following form:

$$Q_{t_i} = \begin{bmatrix} \xi_1 00 \\ 0 \xi_2 0 \\ 00 \xi_3 \end{bmatrix},$$

where $\xi_j = \frac{\theta_j \sigma_j^2}{2k_j}(1 - e^{-k_j \Delta t})^2 + \frac{\sigma_j^2}{k_j}(e^{-k_j \Delta t} - e^{-2k_j \Delta t}) X_j(t_{i-1})$

for $j = 1, 2, 3$.

It is further assumed that the error terms of the measurement (ε_t) and transition equations (u_t) are not correlated.

V The Kalman Filter

Now that we have put the model in equation (4) in state space form, as defined in equations (11) and (12), we can use the Kalman filter algorithm to obtain information about X_t from the observed zero coupon yields. Although the construction of the Kalman filter relies on the normality assumption of the disturbances (both ε_t and u_t are Gaussian noise) and initial state vector, it can calculate the likelihood function by decomposing the prediction error. In the process, it enables the estimation of the parameters by maximum likelihood methods. This is achieved by calculating recursively the distribution of X_t, conditional on the observations at time *t*.

We now describe the filter. At time *t* - 1 we shall have current estimates of the state variables X_{t-1}, the variance Q_{t-1} of X_{t-1}, and the parameters ψ_{t-1}. Starting values of X_0 and Q_0 are provided. There are now three steps involved.

(i) The prediction step in which we find the following:

$X_{t|t-1}$, the forecast of X_t at time *t* - 1.

$Q_{t|t-1}$, the forecast of Q_t at time *t* - 1.

At time *t* we get a new observation, R_t

(ii) The update step. Using R_t, compute estimates of X_t and Q_t

(iii) The parameter estimation step. Using X_t and Q_t, compute an estimate of ψ_t of ψ.

Prediction Step

A forecast of the unobserved state variable X at time *t* is made, based on its value at time *t*-1. These forecasts are, in effect, unbiased conditional estimates.

$$\hat{X}_{t/t-1} = E_{t-1}(X_t) = d + \phi \hat{X}_{t-1} \qquad (13)$$

while the forecast of Q_t, the variance of the measurement error is,

$$Q_{t/t-1} = \phi_{t-1} Q_{t-1} \phi_{t-1}^{'} + Q_{t-1} \qquad (14)$$

For a one-factor CIR model,

$$Q_t = \frac{\theta\sigma^2}{2k}(1 - e^{-k\Delta t})^2 + \frac{\sigma^2}{k}(e^{-k\Delta t} - e^{-2k\Delta t})\hat{X}_{t-1}$$

Update Step

This uses the additional information at time *t*, i.e. R_t to obtain an updated estimator of X_t. When R_t has been observed, the forecast error v_t is

$$v_t = R_t - Z_t \hat{X}_{t|t-1} - d_t$$

The variance F_t of v_t

$$F_t = ZQ_{t|t-1}Z^{'} + H$$

The new estimates X_t and Q_t, in terms of v_t and F_t are

$$\hat{X}_t = X_{t|t-1} + Q_{t|t-1}Z^{'}F_t^{-1}v_t$$

$$Q_t = Q_{t|t-1} - Q_{t|t-1}Z^{'}F_k^{-1}ZQ_{t|t-1}$$

These are the variance minimising conditionally unbiased estimates of X_t and Q_t.

The Log Likelihood Function

The prediction and update steps must be repeated for each discrete time step in the data sample. For the analysis in this paper, I use weekly observations over a period of five years. We use the log likelihood function to estimate the parameter values. Assuming that the prediction errors are normally distributed, the log-likelihood function is given by,

$$\log L(R_1,...,R_n;\psi) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log\left|F_t\right| - \frac{1}{2}v_t^{'}F_t^{-1}v_t \quad \ldots (15)$$

Since the prediction error is Gaussian, equation (15) is the quasi maximum likelihood estimator to best explain the observed values of R_t. Both F_t and v_t depend upon the parameter set given by ψ. Therefore, ψ is chosen so as to maximise the likelihood function $\log L$.

VI Data and Estimation Results

Data description

The data comprises 265 weekly observations of zero-coupon yields for UK and German Treasury bonds from January 6 1999 to January 28, 2003. These observations were sampled every Wednesday to take advantage of high liquidity and avoid beginning and end of week effects. The data sets have a panel data structure with a time dimension and a cross-sectional (maturity) dimension. As zero coupon yields are not observed, cross-sectional data on yield to maturity, time to maturity and coupons for UK and German Treasury bonds, were sourced from *Datastream*. Using this data, zero-coupon yields for UK and German Treasury bonds were estimated using the Nelson and Siegel (1987) model. Eight different maturities that would broadly cover the maturity spectrum of the yield curve are considered; they are 2-, 3-, 5-, 7-, 10-, 15-, 20- and 25-year bonds. Table 1 provides the summary statistics for the estimated zero coupon yields.

Table 1 Summary statistics of zero coupon yields

Maturity	Mean Yield		Standard Deviation	
years	GER	UK	GER	UK
2	4.18	4.89	0.8680	0.8603
3	4.36	4.97	0.7898	0.7826
5	4.61	5.00	0.6721	0.6368
7	4.83	4.97	0.5649	0.5245
10	5.05	4.87	0.5136	0.3539
15	5.12	4.73	0.4952	0.2125
20	5.55	4.60	0.3945	0.2237
25	5.56	4.48	0.3868	0.2416

Figure 2 depicts the evolution of the UK Treasury yield curve over the period of observation. Similarly, Figure 3 shows the evolution of the German Treasury yield curve.

What is immediately discernible from the dynamic behaviour of the UK term structure is the inversion of the yield curve during a phase that would correspond to the year 2000. This reflected the market's fears of worsening economic conditions for the UK, which subsequently eased. In contrast, Germany had positive yield spreads during the same period.

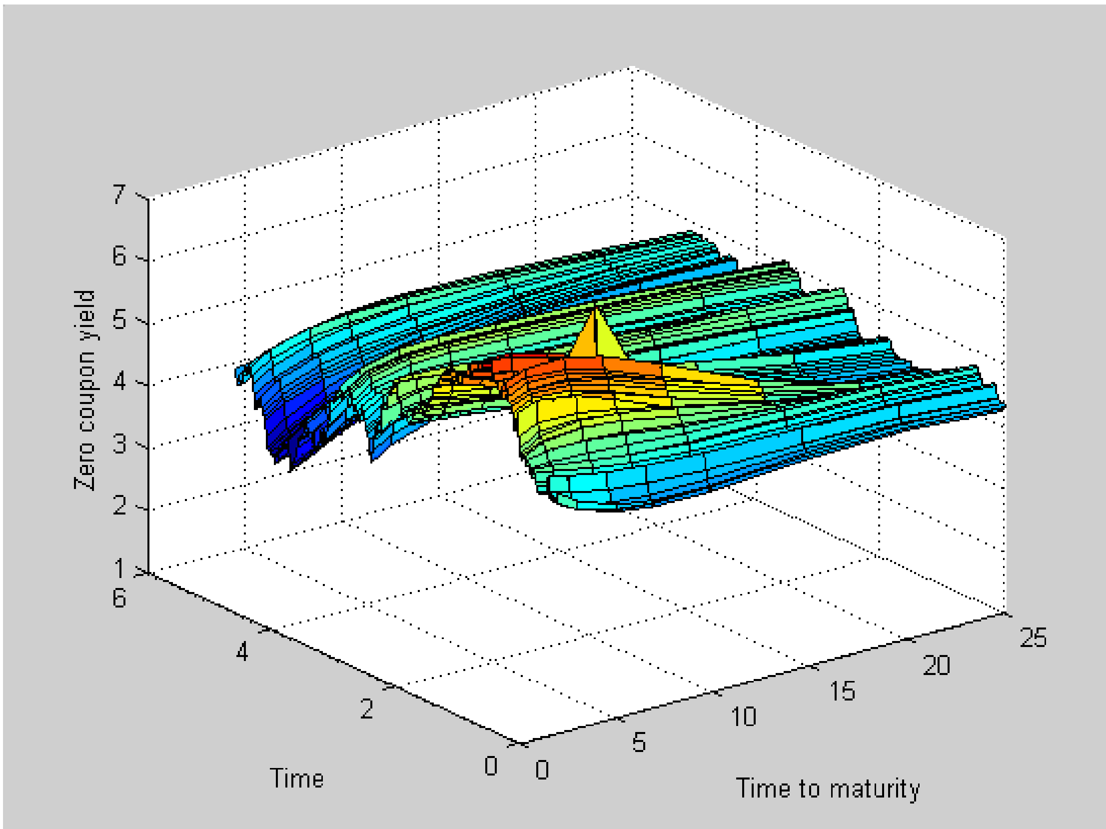

Figure 2. Evolution of the UK Treasury Yield Curve

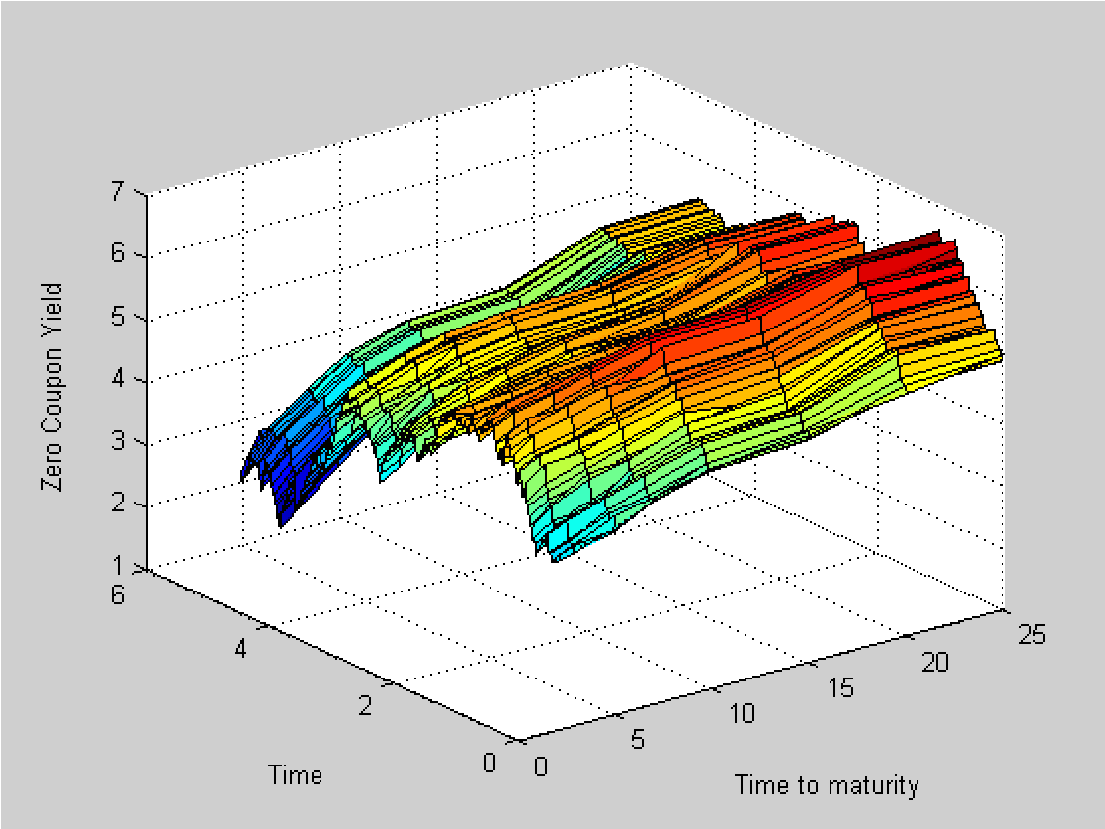

Figure 3. Evolution of the German Treasury Yield Curve

We employ the Kalman filter to estimate the one-factor CIR model using data on the UK and German term structure of interest rates. The objective is to estimate the parameters of the processes that are posited to drive interest rate changes.

Parameter Estimation

The standard errors of the parameter vector $\psi = (\kappa, \theta, \sigma, \lambda, h_1, \cdots, h_8)$ can be computed by using the result shown by White (1982). He showed that the covariance matrix for $\sqrt{n}(\hat{\psi} - \psi)$ converges to

$$\left[E\left(\frac{\partial^2 L}{\partial\psi_i\partial\psi_j}\right)\right]^{-1} E\left(\frac{\partial L}{\partial\psi_i}\frac{\partial L}{\partial\psi_i}\right)\left[E\left(\frac{\partial^2 L}{\partial\psi_i\partial\psi_j}\right)\right]^{-1}$$

where L is the log-likelihood function. The standard errors are given by the diagonals of the above matrix result. Thus for each observation, we determine numerically the partial derivatives of the likelihood with respect to the twelve parameters $\psi = (\kappa, \theta, \sigma, \lambda, h_1, \cdots, h_8)$ evaluated at the maximum likelihood estimate $\hat{\psi}$.

Estimation Results

In keeping with the different dynamics of the term structure observed in the two markets we chose different starting values. For the UK term structure, the initial starting values that have been chosen for the parameters are $\kappa = 0.15$, $\theta = 0.05$, $\sigma = 0.1$, $\lambda = -0.1$. Results of the parameter estimation using the Kalman filter are shown in Table 2. Figures in parenthesis indicate t-values.

Table 2 The Kalman Filter estimates of the one-factor CIR model for UK Treasury bond yields

κ	θ	σ	λ
0.1443	0.0879	0.0801	-0.1176
(3.45)	(3.46)	(3.76)	(2.53)

We obtain significant parameter estimates for all the parameters at the 5% level The significant mean reversion parameter of 0.1443 implies mean reversion in the underlying interest rate. The estimate of 0.1443 indicates a mean half life of 4.8 years which is the expected time for the short rate to return halfway to its long-run average mean, θ. [6] Half-life gives the *slowness* of the mean reversion process and a value of 4.8 years would indicate slow mean reversion for interest rates. But this is comparable to the 4 years obtained by Geyer and Pichler (1999). Accordingly, this process is also characterised by a low but significant volatility estimate (σ = 0.0801). The market price of risk (λ = -0.1176) is negative, a necessary condition for positive risk premia.

In the case of the German term structure, the initial starting values that have been chosen for the parameters are $\kappa = 0.15$, $\theta = 0.04$, $\sigma = 0.05$, $\lambda = -0.1$. Results of the parameter estimation using the Kalman filter are shown in Table 3. Figures in parenthesis indicate t-values.

Table 3 The Kalman Filter estimates of the one-factor CIR model for German Treasury bond yields

κ	θ	σ	λ
0.1579	0.0646	0.0556	-0.00095
(20.83)	(15.1)	(2.37)	(0.12)

[6] The half life is given by $e^{-kt} = 0.5$. This implies $t = -\ln(0.5)/k$

We obtain significant parameter estimates for all the parameters except the market price of risk. This would suggest that the variable has not been priced by the market. In accordance with the lower level of short-term yields for German Treasury bonds, the long-term mean parameter is 6.46 per cent as compared to 8.79 per cent for the UK Treasury. The mean reversion of 0.1579 implies a mean half-life of 4.38 years and this is somewhat smaller in magnitude as compared to that obtained for the UK term structure. However, the volatility parameter given by 0.0556 is significantly smaller than that obtained for the UK term structure The market price of risk parameter is estimated with no significant precision and this could be indicative of its correlation with the long-term mean parameter.

VII Conclusion

In this paper a single-factor CIR model has been estimated for the UK and German term structure for the period January, 1999 to January, 2004. Modelling continuous time term structure models start out with the specification of a time series process for the instantaneous spot interest. The no-arbitrage condition then permits the derivation of a bond pricing formula whereby the bond price is a function of the unobserved instantaneous spot rate and the model's parameters. These parameters are the long-run mean, the speed of adjustment towards the long-run mean, volatility of the short-term interest rate and the market price of risk. In this paper the model is estimated for a single factor using a quasi maximum likelihood approach based on the Kalman filter. The Kalman filter algorithm uses observable data on bonds to extract values for the unobserved state variables. It combines both the cross section and time series information in the term structure.

The yields on zero-coupon bonds are used as inputs for the estimation process. In this paper, the empirical analysis is based on weekly observations of UK and German Treasury zero coupon bonds over the period January 1999 to January 2004. Eight maturities have been chosen that span the yield curve from 2 years to 25 years and are expected to incorporate influences on the short, medium and long end of the term structure. The parameters of the model and their standard errors are estimated.

Results of the empirical analysis show that the unobserved instantaneous interest rate exhibits mean reverting behaviour in both the UK and German term structure. However, the mean reversion of the interest rate process has been relatively slower in the UK as compared to Germany since the introduction of the euro. Accordingly, the volatility component, which shocks the process at each step in time is also higher in the UK as compared to Germany. The results indicate that the one-factor CIR model provides a good representation of the UK Gilt-Edged market. But its inability to meaningfully account for the market price of risk has impinged on its efficacy in capturing the dynamics of the German term structure.

References

Babbs, S.H. and K.B. Nowman (1999). "Kalman Filtering of Generalised Vasicek Term Structure Models" . *Journal of Financial and Quantitative Analysis,* pp 115-130.

Bolder, D.J. (2001), "Affine Term-Structure Models: Theory and Implementation." *Bank of Canada Working Paper 2001-15*

Chan, K.C., G.A. Karoly, F.A. Longstaff and A.B. Sanders (1992). "An Empirical Comparison of Alternative Models of the Short-Term Interest." *Journal of Finance,* 47 pp. 183-193.

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

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

Duan J.C. and J.G. Simonato (1998). "Estimating and testing exponential affine term structure models by the Kalman Filter." *Review of Quantitative Finance and Accounting*.

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

Geyer, A.L.J., and Pichler, S., (1999). "A State-Space Approach to Estimate and Test Multifactor Cox-Ingersoll-Ross Models of the Term Structure of Interest Rates," *Journal of Financial Research,* Vol 22(1), 107-130.

Harvey, A.C. (1989). "Forecasting, Structural Time Series Models and the Kalman Filter." (Cambridge University Press)

Hull, J.C., (2000). "Options, Futures and other Derivatives." Pearson Education Inc.

James, J., and N. Webber., (2002). "Interest Rate Modelling." John Wiley & Sons

Litterman, R. and Scheinkman, J. (1991). "Common factors affecting bond returns," *Journal of Fixed Income* (June): 54-61

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

Lund, J. (1997). "Econometric Analysis of Continuous-Time Arbitrage-Free models of the Term Structure of Interest Rates," Department of Finance, The Aarhus School of Business

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

Nelson, C.R., and Siegel, A.F., (1987), "Parsimonious modelling of yield curves," Journal of Business Vol. 60: 474-489.

Richard, S.F. (1978): "An Arbitrage Model of the Term Structure of Interest Rates," *Journal of Financial Economics,* 6, 33-37.

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