Working Paper Series

24/2006

Modelling the long-term dynamics of yield curves

Medova, E.A., Rietbergen, M.I., Villaverde, M. and Yong, Y.S.

CAMBRIDGE
Judge Business School

Author contact details are as follows:

Dr E A Medova
Centre for Financial Research
Judge Business School
University of Cambridge
eam28@cam.ac.uk

M I Rietbergen
Morgan Stanley
25 Cabot Square
London E14 4QA
muriel.rietbergen@morganstanley.com

Dr M Villaverde
Marshall Wace Asset Management
The Adelphi, 13th fl
1/11 John Adam Street
London WC2N 6HT
m.villaverde.mwam.com

Y S Yong
Centre for Financial Research
Judge Business School
University of Cambridge
ysy20@cam.ac.uk

Please address enquiries about the series to:

Research Support Manager
Judge Business School
Trumpington Street
Cambridge CB2 1AG, UK
Tel: 01223 760546 Fax: 01223 339701
E-mail: research-support@jbs.cam.ac.uk

Modelling the Long-Term Dynamics of Yield Curves

E.A. Medova, M.I. Rietbergen, M. Villaverde and Y.S. Yong

Dr E.A. Medova[1]
Centre for Financial Research
Judge Business School
University of Cambridge
&
Cambridge Systems Associates Ltd.

Dr M.I. Rietbergen
Centre for Financial Research
Judge Business School
University of Cambridge
&
Cambridge Systems Associates Ltd.

Dr M. Villaverde
Marshall Wace Asset Management
The Adelphi
13th Floor
1/11 John Adam Street
London WC2N 6HT

Y.S. Yong
Centre for Financial Research
Judge Business School
University of Cambridge
&
Cambridge Systems Associates Ltd.

Abstract

This paper presents a three-factor model of the term structure of interest rates, which is Markov and time-homogeneous. We provide a thorough analysis of the estimation procedure using the Kalman filter on EU swap yield data from 1997 to 2002. The model allows for a closed-form bond price formula and can capture the salient features of the whole term structure in forward simulations. These features make it particularly useful for applications in long-term asset pricing, risk management and portfolio optimization.

JEL Classification: G10, G12.

Keywords: Multifactor Term Structure Model; Kalman Filter; Simulations.

[1] Dr E.A. Medova, Cambridge Systems Associates, 5-7 Portugal Place, Cambridge CB5 8AF, United Kingdom, Tel: +44 1223 557640, Fax: +44 1223 557641, Email: eam28@cam.ac.uk

1. Introduction

The literature in the area of interest rate modelling is extensive. Traditional term structure models, such as Vasicek (1977) and Cox, Ingersoll and Ross (1985) take as given a short rate process and its market price of risk. Ho and Lee (1986) and Heath, Jarrow and Morton (1992) followed this work with a new approach to interest rate modelling in which they fit the initial term structure exactly. More recent studies are given by Duffie and Kan (1996), Jamshidian (1997) and the book by James and Webber (2000). Development of these models however has been driven primarily by the need for models to price and hedge relatively short-term derivatives. One of the most recent papers on this topic is by Andersen *et al.* (2004), who consider a three-factor model with stochastic volatility, mean drift and jumps, but again their focus is on modelling the short-term interest rate (3-month U.S. T-bill).

Until recently, little research has been conducted on the development of models that satisfy realistic dynamics in the long-term. With the move in recent years from defined benefit plans to defined contribution plans in the pension world, we now see a renewed interest in insurance products with long-term guarantees. However, unlike during the 1970s, when the option element introduced by guarantees was often ignored in the liability pricing, historically low interest rates in recent years have emphasized the importance of accurately valuing long-term guarantees (Wilkie *et al.*, 2004). Banks also face new pricing challenges due to the increased demand for long-maturity derivatives (e.g. 60-year swaps) and therefore require good long-term interest rate models. Moreover, accounting practices are moving more and more towards using the fair value of assets and liabilities in balance sheets which also requires long-term models.

In Figure 1 we plot the development over time of short- and long-term interest rates in the Eurozone for the period 1997-2002. Figure 2 plots the weekly standard deviations of the yields over the same period. As the short-term and long-term rates are not perfectly correlated, the data are clearly inconsistent with the use of a one-factor time-homogeneous model. Chan *et al.* (1992) demonstrate the empirical difficulties of one-factor continuous-time specifications within the Vasicek and Cox, Ingersoll and Ross class of models using the generalized methods of moments.

Litterman and Scheinkman (1991) find that 96% of the variability of the excess returns of any zero-coupon bond can be explained by three factors: the *level*, *steepness* and *curvature*. They also point out that the 'correct model' of the term structure may involve unobservable factors. For instance, it is widely believed that changes in the Federal Reserve policy are a major source of changes in the shape of the US yield curve. Even though the Federal Reserve policy itself is observable, it is not clear how to measure its effect on the yield curve. In fact, Litterman and Scheinkman (1991) themselves used unobservable factors in their approach by applying principal component analysis.

Most term structure models such as Ho and Lee (1986), Hull and White (1990) and Heath, Jarrow and Morton (1989) are specified using the risk-neutral measure. This makes them appropriate for relative-pricing applications, but inappropriate for forward simulations, which needs to take place under the real-world measure. An exception is Rebonato *et al.* (2005) who focus on the yield curve evolution under the real-world measure and present a semi-parametric method to explain the yield curve evolution.

In this paper we focus on a term structure model with the following characteristics:

- The model is set in a continuous-time framework. This allows implementation in discrete time with any length of time step, Δt, without the need to construct a new model each time we change Δt. This is an important requirement for the flexibility of forward simulations.
- Interest-rate dynamics are consistent with what we observe in historical data.
- The model has a closed-form solution for bond pricing, permitting straightforward analytical calculation in simulation.
- The short rate is mean-reverting.
- The model permits a tractable method of estimation and calibration.
- The model is flexible enough to give rise to a range of different yield curve shapes and dynamics (steepening, flattening, yield curve inversion, etc.).

The remainder of the paper is structured as follows. In Section 2, a three-factor term structure model is introduced and a closed-form solution for the bond price derived. Section 3 discusses the state-space formulation of the models and the estimation of the parameters using the Kalman filter. The data and empirical analysis, focusing on fitting the data as well as on the simulation potential of the model, are presented in Section 4. Finally Section 5 concludes.

2. Three-Factor Term Structure Model

The term structure model presented in this paper is driven by three factors as proposed by Litterman and Scheinkman (1991) and can be viewed as an extension to the generalized Vasicek model presented by Langetieg (1980). The first two factors **X** and **Y** satisfy the standard Vasicek stochastic differential equations with mean reversion levels of $\frac{\mu_X}{\lambda_X}$ and $\frac{\mu_Y}{\lambda_Y}$ respectively. The difference is in the way the short rate **R** is modelled. In this case the mean reversion level is stochastic rather than deterministic and depends on the level of the other two factors **X** and **Y** driving the model.

Starting from the formulation of the model under the risk neutral measure, *Q*, we have the following three stochastic differential equations (SDEs):

$$d\mathbf{X}_t = (\mu_X - \lambda_X X_t)dt + \sigma_X d\tilde{\mathbf{W}}_t^X \tag{1}$$

$$d\mathbf{Y}_t = (\mu_Y - \lambda_Y Y_t)dt + \sigma_Y d\tilde{\mathbf{W}}_t^Y \tag{2}$$

$$d\mathbf{R}_t = k(X_t + Y_t - R_t)dt + \sigma_R d\tilde{\mathbf{W}}_t^R, \tag{3}$$

where the $d\tilde{\mathbf{W}}$ terms are correlated. Factoring the covariance matrix of the $d\tilde{\mathbf{W}}$ terms using a Cholesky decomposition results in the following formulation:

$$d\mathbf{X}_t = (\mu_X - \lambda_X X_t)dt + \sum_{i=1}^{3} \sigma_{X_i} d\mathbf{Z}_t^i \tag{4}$$

$$d\mathbf{Y}_t = (\mu_Y - \lambda_Y Y_t)dt + \sum_{i=1}^{3} \sigma_{Y_i} d\mathbf{Z}_t^i \tag{5}$$

$$d\mathbf{R}_t = k(X_t + Y_t - R_t)dt + \sum_{i=1}^{3} \sigma_{R_i} d\mathbf{Z}_t^i, \tag{6}$$

where the $d\mathbf{Z}$ terms are uncorrelated.

To obtain the zero-coupon bond price we first solve the SDE for **X**. Integrating

$\int d(e^{\lambda_X u}\mathbf{X_u})$ gives

$$\mathbf{X}_s = \frac{\mu_X}{\lambda_X} + \left(X_t - \frac{\mu_X}{\lambda_X} \right) e^{\lambda_X (t-s)} + \sum_{i=1}^{3} \sigma_{X_i} \int_t^s e^{\lambda_X (u-s)} d\mathbf{Z}_u^i. \tag{7}$$

Solving for **Y** in the same manner gives

$$\mathbf{Y}_s = \frac{\mu_Y}{\lambda_Y} + \left(Y_t - \frac{\mu_Y}{\lambda_Y} \right) e^{\lambda_Y (t-s)} + \sum_{i=1}^{3} \sigma_{Y_i} \int_t^s e^{\lambda_Y (u-s)} d\mathbf{Z}_u^i \tag{8}$$

Integrating $\int d(e^{ku}\mathbf{R}_u)$ and substituting the solution for **X** and **Y** above gives the following solution for **R**

$$\begin{aligned}
\mathbf{R}_s = {} & e^{k(t-s)} R_t + \frac{\mu_X}{\lambda_X}(1 - e^{k(t-s)}) + \frac{k}{k - \lambda_X}\left(X_t - \frac{\mu_X}{\lambda_X} \right)(e^{\lambda_X (t-s)} - e^{k(t-s)}) + \\
& k \sum_{i=1}^{3} \frac{\sigma_{X_i}}{k - \lambda_X} \int_t^s (e^{\lambda_X (u-s)} - e^{k(u-s)}) d\mathbf{Z}_u^i + \\
& \frac{\mu_Y}{\lambda_Y}(1 - e^{k(t-s)}) + \frac{k}{k - \lambda_Y}\left(Y_t - \frac{\mu_Y}{\lambda_Y} \right)(e^{\lambda_Y (t-s)} - e^{k(t-s)}) + \\
& k \sum_{i=1}^{3} \frac{\sigma_{Y_i}}{k - \lambda_Y} \int_t^s (e^{\lambda_Y (u-s)} - e^{k(u-s)}) d\mathbf{Z}_u^i + \\
& \sum_{i=1}^{3} \sigma_{R_i} \int_t^s e^{k(u-s)} d\mathbf{Z}_u^i.
\end{aligned} \tag{9}$$

Let $P(t,T)$ denote the price of a zero-coupon bond at time *t* paying 1 at time *T*, then

$$P(t,T) = \mathbb{E}_t^Q\left\{\exp\left(-\int_t^T \mathbf{R}_s ds\right)\right\}, \tag{10}$$

where $\mathbb{E}_t^Q$ denotes the expectation under the risk neutral measure Q conditional on the information at time t. As $\mathbf{R}_s$ is normally distributed in our model we can use the moment generating function for the normal distribution to rewrite (10) as

$$P(t,T) = \exp\left\{\mathbb{E}_t^Q\left(-\int_t^T \mathbf{R}_s ds\right) + \frac{1}{2}\mathrm{var}_t^Q\left(-\int_t^T \mathbf{R}_s ds\right)\right\}, \tag{11}$$

where var_t^Q denotes the conditional variance under Q. Integrating the solution (9) for **R** and taking the expectation and variance of the result gives expressions for the two terms in (11) involving the parameters

$$\begin{aligned} m_{X_i} &= -\frac{k\sigma_{X_i}}{\lambda_X(k-\lambda_X)} \\ m_{Y_i} &= -\frac{k\sigma_{Y_i}}{\lambda_Y(k-\lambda_Y)} \\ n_i &= \frac{\sigma_{X_i}}{k-\lambda_X} + \frac{\sigma_{Y_i}}{k-\lambda_Y} - \frac{\sigma_{R_i}}{k} \\ p_i &= -(m_{X_i} + m_{Y_i} + n_i). \end{aligned} \tag{12}$$

Hence

$$P(t,T) = \exp\{-A(t,T)R_t - B(t,T)X_t - C(t,T)Y_t - D(t,T)\} \tag{13}$$

and

$$y_{t,T} = \frac{A(t,T)R_t + B(t,T)X_t + C(t,T)Y_t + D(t,T)}{T-t}, \tag{14}$$

where

$$A(t,T)=\frac{1}{k}(1-e^{-k(T-t)}) \tag{15}$$

$$B(t,T)=\frac{k}{k-\lambda_X}\left\{\frac{1}{\lambda_X}(1-e^{-\lambda_X(T-t)})-\frac{1}{k}(1-e^{-k(T-t)})\right\} \tag{16}$$

$$C(t,T)=\frac{k}{k-\lambda_Y}\left\{\frac{1}{\lambda_Y}(1-e^{-\lambda_Y(T-t)})-\frac{1}{k}(1-e^{-k(T-t)})\right\} \tag{17}$$

$$\begin{aligned}D(t,T)=&\left(T-t-\frac{1}{k}(1-e^{-kT})\right)\left(\frac{\mu_X}{\lambda_X}+\frac{\mu_Y}{\lambda_Y}\right)-\frac{\mu_X}{\lambda_X}B(t,T)-\frac{\mu_Y}{\lambda_Y}C(t,T)-\\&\frac{1}{2}\sum_{i=1}^{3}\left\{\frac{m_{X_i}^2}{2\lambda_X}(1-e^{-2\lambda_X(T-t)})+\frac{m_{Y_i}^2}{2\lambda_Y}(1-e^{-2\lambda_Y(T-t)})+\right.\\&\frac{n_i^2}{2k}(1-e^{-2k(T-t)})+p_i^2(T-t)+\frac{2m_{X_i}m_{Y_i}}{\lambda_X+\lambda_Y}(1-e^{-(\lambda_X+\lambda_Y)(T-t)})+\\&\frac{2m_{X_i}n_i}{\lambda_X+k}(1-e^{-(\lambda_X+k)(T-t)})+\frac{2m_{X_i}p_i}{\lambda_X}(1-e^{-\lambda_X(T-t)})+\\&\frac{2m_{Y_i}n_i}{\lambda_Y+k}(1-e^{-(\lambda_Y+k)(T-t)})+\frac{2m_{Y_i}p_i}{\lambda_Y}(1-e^{-\lambda_Y(T-t)})+\\&\left.\frac{2n_ip_i}{k}(1-e^{-k(T-t)})\right\}.\end{aligned} \tag{18}$$

Bond pricing can be achieved under the *risk-neutral* measure *Q*. However, for the model to be used for forward simulations, we will have to adjust the set of stochastic differential equations so that we capture the model dynamics under the *real-world* measure *P*. We therefore have to model the market prices of risk which take us from the risk-neutral measure *Q* to the real-world measure *P*.

Under the real-world measure *P*, we must adjust the drift term by adding the *risk premium* which is given by the *market price of risk* γ times the quantity of risk. The effect of this is a change in the long-term mean, i.e. for the factor $\mathbf{X}$ the long-term mean becomes $\frac{\mu_X + \gamma_X \sigma_X}{\lambda_X}$. It is generally assumed that in a Gaussian world the quantity of risk is given by the *volatility* of each factor.

It can be shown that the market prices of risk will be independent of the time to maturity of the bond and of the factor being modelled. However, time homogeneity is usually just assumed, as in (Vasicek, 1977). The set of processes under the real world measure are thus:

$$d\mathbf{X}_t = (\mu_X - \lambda_X X_t + \gamma_X \sigma_X)dt + \sigma_X d\mathbf{W}_t^X \tag{19}$$

$$d\mathbf{Y}_t = (\mu_Y - \lambda_Y Y_t + \gamma_Y \sigma_Y)dt + \sigma_Y d\mathbf{W}_t^Y \tag{20}$$

$$d\mathbf{R}_t = \{k(X_t + Y_t - R_t) + \gamma_R \sigma_R\}dt + \sigma_R d\mathbf{W}_t^R, \tag{21}$$

where all three factors contain a market price of risk γ in volatility units.

3. Kalman Filter

There is always going to be a trade-off between the richness of the econometric representation of the state variables and the computational burden of estimation and pricing. Affine term structure models yield essentially closed-form expressions for zero-coupon bond prices (Duffie and Kan, 1996) which greatly facilitate pricing and econometric implementation.

The focus is on trying to fit the yields and think of the factors as unobserved input variables. To handle the unobservable state variables we formulate the model in state-space form, a detailed description of which can be found in Harvey (1993), and use the Kalman filter recursively to estimate the model parameters.

The state-space approach adopted in this paper simultaneously integrates time-series and cross-sectional aspects of the yield curve model. Moreover, it allows the identification of the market prices of interest rate risk. In the state-space model there is a *transition equation* for the latent factors and a *measurement equation* for the yields on an arbitrary number of maturities.

Some examples of the growing literature that estimates term structure models using panel data is given by Babbs and Nowman (1999), Chen and Scott (1993), De Jong (2000), De Jong and Santa-Clara (1999) and Geyer and Pichler (1997). Most of these papers analyze multi-factor versions of the Cox-Ingersoll-Ross (CIR) model using mutually independent factors. De Jong (2000) extends this approach to the more general class of affine models proposed by Duffie and Kan (1996).

State-space form (SSF) is a powerful tool which allows us to handle a wide range of time series models. The general state-space form applies to multivariate time series. The *N* observable variables at time *t*, $\mathbf{y}_t$, are related to a vector $\boldsymbol{\alpha}_t$ known as the *state vector* via a *measurement equation*

$$\mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t + \mathbf{d} + \boldsymbol{\varepsilon}_t \qquad t = 1, \ldots, T, \qquad (22)$$

where **Z** is an $N \times m$ matrix and **d** and $\boldsymbol{\varepsilon}_t$ are $N \times 1$ vectors, where the error term is assumed to consist of serially uncorrelated disturbances with mean zero and covariance matrix **H**, i.e.

$$\mathbb{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0} \qquad \text{var}(\boldsymbol{\varepsilon}_t) = \mathbf{H}. \tag{23}$$

In general **Z**, **d** and **H** may depend on *t*.

Even though the elements of $\boldsymbol{\alpha}_t$ tend to be unobservable, they are known to follow the following first-order Markov process, which is known as the *transition equation*

$$\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \mathbf{c} + \mathbf{S}\boldsymbol{\eta}_t \qquad t = 1, \ldots, T, \tag{24}$$

where **T** is an $m \times m$ matrix, **c** an $m \times 1$ vector, **S** an $m \times g$ matrix and $\boldsymbol{\eta}_t$ a $g \times 1$ vector of serially uncorrelated disturbances with mean zero and covariance matrix **Q**, that is

$$\mathbb{E}(\boldsymbol{\eta}_t) = \mathbf{0} \qquad \text{var}(\boldsymbol{\eta}_t) = \mathbf{Q}. \tag{25}$$

Again, in general **T**, **c** and **S** may depend on *t*, however we will be dealing with the *time-homogeneous* case.

Two further assumptions will be required to complete the state-space formulation:

- The initial state vector $\boldsymbol{\alpha}_0$ has a mean of $\mathbf{a}_0$ and a covariance matrix $\mathbf{P}_0$, that is

$$\mathbb{E}(\boldsymbol{\alpha}_0) = \mathbf{a}_0 \qquad \text{var}(\boldsymbol{\alpha}_0) = \mathbf{P}_0. \tag{26}$$

- The disturbance terms $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\eta}_t$ are uncorrelated with each other in all time periods and uncorrelated with the initial state, that is

$$\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\eta}_s') = \mathbf{0} \qquad \text{for all } s, t = 1, \ldots, T \tag{27}$$

and

$$\mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\alpha}_0') = \mathbf{0} \qquad \mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\alpha}_0') = \mathbf{0} \qquad t = 1, \ldots, T. \tag{28}$$

The important concept behind the state space formulation is the separation of the noise driving the system dynamics $\boldsymbol{\eta}_t$ and the observational noise $\boldsymbol{\varepsilon}_t$.

The *Kalman filter* is applied recursively in order to compute the optimal estimator of the state vector at time *t* given all the information currently available, which consists of the observations up to and including $\mathbf{y}_t$. Assuming a Gaussian state space, the disturbances and the initial state vector will be normally distributed, and hence the likelihood function can be calculated using prediction error decomposition.

In a state-space model the system matrices will usually depend on a set of unknown parameters, in our case 14, referred to as *hyper-parameters* and defined in Table 1 below. Using the Kalman filter to construct the likelihood function and then maximizing it using a suitable numerical optimization procedure, we can carry out maximum likelihood estimation of the hyper-parameters. The joint probability of a set of T observations can be expressed in terms of conditional distributions. For a multivariate normal we have

$$L(\mathbf{y};\boldsymbol{\varphi}) = \prod_{t=1}^{T} p(\mathbf{y}_t \left| \mathbf{Y}_{t-1}\right.), \qquad (29)$$

where $p(\mathbf{y}_t \mid \mathbf{Y}_{t-1})$ is the distribution of $\mathbf{y}_t$ conditional on the information at time $t-1$, i.e. $\mathbf{Y}_{t-1} = (y_{t-1}, y_{t-2}, ..., y_1)'$. Since we have a Gaussian model we can write the log-likelihood function in *prediction error decomposition form* as

$$\log L(\boldsymbol{\varphi}) = -\frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log\left|\mathbf{F}_t\right| - \frac{1}{2}\sum_{t=1}^{T}\mathbf{v}_t'\mathbf{F}_t^{-1}\mathbf{v}_t, \qquad (30)$$

where $\mathbf{F}_t$ is estimated by the covariance matrix obtained from the Kalman filter as

$$\mathbf{F}_t = \mathbf{Z}\mathbf{P}_{t|t-1}\mathbf{Z}' + \mathbf{H} \qquad (31)$$

and $\mathbf{v}_t$ is the vector of *prediction errors* given by

$$\mathbf{v}_t = \mathbf{y}_t - \tilde{\mathbf{y}}_{t|t-1} = \mathbf{Z}(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t|t-1}) + \boldsymbol{\varepsilon}_t. \qquad (32)$$

Together with the following two equations, (31) and (32) form the *measurement update equations*

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{Z}'\mathbf{F}_t^{-1}\mathbf{v}_t \tag{33}$$

$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{Z}'\mathbf{F}_t^{-1}\mathbf{P}_{t|t-1}. \tag{34}$$

So first we specify starting values for the parameters. With these starting values we run the Kalman filter to obtain estimated yields and a time series for the unobserved state variables. Next, the parameters are estimated by maximizing the log-likelihood using the variable path estimates as observations. The optimized parameter values are then used as the starting values for the next iteration of the Kalman filter. This loop continues until we obtain the optimal parameter estimates (*cf.* Dempster *et al.*, 1977). The calibration code is implemented in C++ and the optimization is performed using a combination of global (Direct, see Jones *et al.*, 1993) and local (approximate) conjugate direction (Powell, 1964) numerical algorithms.

The starting values for the Kalman filter are given by the mean and the covariance of the unconditional distribution of the stationary state vector. The state vector is stationary if $\mathbf{c}$ and $\mathbf{T}$ are time invariant and $|\lambda(\mathbf{T})| < 1$, where $\lambda(\mathbf{T})$ is the leading eigenvalue of $\mathbf{T}$. In this case the mean $\mathbf{a}_0$ is given by the unique solution to

$$\mathbf{a}_0 = \mathbf{T}\mathbf{a}_0 + \mathbf{c} \qquad \Rightarrow \qquad \mathbf{a}_0 = (\mathbf{I} - \mathbf{T})^{-1}\mathbf{c} \tag{35}$$

and the covariance matrix $\mathbf{P}_0$ will be given by the unique solution to the *Riccati equation*

$$\mathbf{P}_0 = \mathbf{T}\mathbf{P}_0\mathbf{T}' + \mathbf{S}\mathbf{Q}\mathbf{S}' \qquad \Rightarrow \qquad vec(\mathbf{P}_0) = (\mathbf{I} - \mathbf{T} \otimes \mathbf{T})^{-1} vec(\mathbf{S}\mathbf{Q}\mathbf{S}'). \tag{36}$$

In our case the observable variables are given by swap yields of different maturities, and are related to the vector of unobservable state variables $\left(\mathbf{X},\mathbf{Y},\mathbf{R}\right)$ via the measurement equation. The measurement equation is obtained by using (14) and adding serially and cross-sectionally uncorrelated disturbances with mean zero to take into account non-simultaneity of the observations, errors in the data, etc. The unobservable state variables are generated via the transition equations, which in our case are given by the discretized versions of (1), (2) and (3), using Euler's first order approximation, i.e.

$$\mathbf{X}_{t+\Delta t} = X_t + (\mu_X - \lambda_X X_t + \gamma_X \sigma_X)\Delta t + \sigma_X \sqrt{\Delta t}\,\boldsymbol{\eta}_{t,X} \tag{37}$$

$$\mathbf{Y}_{t+\Delta t} = Y_t + (\mu_Y - \lambda_Y Y_t + \gamma_Y \sigma_Y)\Delta t + \sigma_Y \sqrt{\Delta t}\,\boldsymbol{\eta}_{t,Y} \tag{38}$$

$$\mathbf{R}_{t+\Delta t} = R_t + (k(X_t + Y_t - R_t) + \gamma_R \sigma_R)\Delta t + \sigma_R \sqrt{\Delta t}\,\boldsymbol{\eta}_{t,R}. \tag{39}$$

In matrix form the transition equations can now be written as

$$\begin{pmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \\ \mathbf{R}_t \end{pmatrix} = \mathbf{T} \begin{pmatrix} X_{t-\Delta t} \\ Y_{t-\Delta t} \\ R_{t-\Delta t} \end{pmatrix} + \mathbf{c} + \mathbf{S}\boldsymbol{\eta}_t, \tag{40}$$

where

$$\mathbf{T} = \begin{pmatrix} 1-\lambda_X \Delta t & 0 & 0 \\ 0 & 1-\lambda_Y \Delta t & 0 \\ k\Delta t & k\Delta t & 1-k\Delta t \end{pmatrix} \tag{41}$$

$$\mathbf{c} = \begin{pmatrix} (\mu_X + \gamma_X \sigma_X)\Delta t \\ (\mu_Y + \gamma_Y \sigma_Y)\Delta t \\ \gamma_R \sigma_R \Delta t \end{pmatrix} \tag{42}$$

$$\mathbf{S} = \begin{pmatrix} \sigma_X \sqrt{\Delta t} & 0 & 0 \\ 0 & \sigma_Y \sqrt{\Delta t} & 0 \\ 0 & 0 & \sigma_R \sqrt{\Delta t} \end{pmatrix} \tag{43}$$

and $\boldsymbol{\eta}_t$ is a vector with serially uncorrelated disturbances satisfying

$$\mathbb{E}(\boldsymbol{\eta}_t) = \mathbf{0} \qquad \mathrm{var}(\boldsymbol{\eta}_t) = \begin{pmatrix} 1 & \rho_{XY} & \rho_{XR} \\ \rho_{XY} & 1 & \rho_{YR} \\ \rho_{XR} & \rho_{YR} & 1 \end{pmatrix}. \tag{44}$$

In the current literature, several approaches have been adopted to estimate the covariance matrix of the measurement errors. For example, De Jong and Santa Clara (1999) used a spherical covariance matrix, $\mathbf{H} = h\mathbf{I}$, whereas Babbs and Nowman (1999) use a diagonal matrix. De Jong (2000) uses a full covariance matrix.

We adopted a diagonal covariance matrix approach, optimizing the measurement errors using group one-at-a-time search with two groups: in the first group the model parameters were optimized followed by the optimization of the measurement errors in the second group. This process is repeated until convergence. This method is preferred over the full optimization with 14 model parameters and 16 measurement errors due to the scale of the optimization problem in the latter. Even though the full covariance matrix is to be highly preferred, it is avoided in our case since using yields of 16 different maturities would result in 136 noise parameters to be estimated.

4. Estimation and Simulation Results

For our empirical analysis yields on ordinary, fixed-for-floating rate Euro swap contracts are used. Dai and Singleton (2000) point out that these yields are preferable for analysis for the following reasons. The swap markets provide ‘constant maturity’ yield data, whereas in the Treasury market the maturities of ‘constant maturity’ yields are only approximately constant or the data represent interpolated series. Additionally, the on-the-run treasuries that are often used in empirical studies are typically on ‘special’ in the repo market. So, strictly speaking, the Treasury data should be adjusted for repo specials prior to any empirical analysis. Unfortunately, the requisite data for making these adjustments are not readily available, and, consequently, such adjustments are rarely made.

For estimation and calibration purposes, we used weekly Euro swap data for the period June 1997 to December 2002 (a total of 292 time points) of 16 different yields with maturities equal to 1, 3 and 6 months and 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20 and 30 years. The length of the sample period was determined in part by the unavailability of reliable long-term swap data for years prior to 1997.

The estimation results are presented in Table 1. All parameter estimates have plausible values and all are statistically significant, unlike the estimates found by

Babbs and Nowman (1999), who looked at Kalman filtering generalized Vasicek models. However they only used yields of eight different maturities and Geyer and Pichler (1999) show that a large number of maturities is important to improve the precision of the parameter estimates.

Table 2 provides the estimated standard deviations $\sqrt{h_i}$ of the measurement errors, where h_i is the ith diagonal element of the covariance matrix **H**. In particular, these standard deviations range from less than 1 basis point for the seven-year yield to 24 basis points for the thirty-year rate. These measurement errors compare in magnitude to those in Babbs and Nowman (1999) and compare very favourably to studies by, for example Chen and Scott (1993) and Geyer and Pichler (1996), who both estimate the multifactor Cox-Ingersoll-Ross model on U.S. data.
Similar to Geyer and Pichler (1996), the error standard deviations exhibit a distinct U-shaped pattern as depicted in Figure 4. A possible explanation for this might be that the observed data for the medium range are highly correlated and therefore easier to fit. It also indicates that using the one-month yield as a proxy for the short rate is likely to give rise to serious problems.

Like the Babbs and Nowman (1999) paper, we also look at the factor loadings of this three-factor model as a function of maturity to determine the nature of the factors calculated by the Kalman filter. As factor loadings correspond to orthogonal Brownian motions, rather than correlated innovations, we use a Cholesky decomposition as described in Section 2 to transform the stochastic differential equations. The curve for each factor represents the change in yield caused by a shock from that factor of one standard deviation ($dy_{t,T}/dZ_i$). For comparison with Babbs and Nowman (1999), we also impose the following three additional restrictions: the second factor has approximately zero impact on the term structure at the five-year maturity and the third factor loading disappears at around two and twelve years. This gives a set of nine equations in nine unknowns.

Figure 5 plots the factor loadings for the three-factor model. Whereas Babbs and Nowman found that their third factor loading had a negligible effect, we find all three factors have a significant impact on the yields of all maturities. We also find that the range of the impact of the three factors on the yields is similar to that found by Litterman and Scheinkman (1991).

4.1 Simulation Results

One of the objectives of this paper is to propose a term structure model that is tractable in forward simulations but can still capture the salient features of the yield curve. To test our model, we performed a backtest over 2003. Using the historical 52 weekly data points for the yields over 2003, we calculated the mean level and the weekly standard deviation for each of the sixteen maturities. We then simulated forward from January 2003 to January 2004 using the parameter estimates given in Table 1. In total 500 scenarios were generated and for each the mean and standard deviation for each of the sixteen maturities was calculated. Averaging over all scenarios finally gives an average mean and standard deviation for the simulated yields.

Figure 6 plots the mean level of the yields for both the historical and the simulated data and Figure 7 plots the standard deviation. As can be observed from Figure 6 the two sets of means closely match each other. Figure 7 shows that the simulated standard deviations slightly over-estimate the historical ones. However, as Figure 8 shows, yields were more stable in 2003 relative to 1997-2002, which would explain this discrepancy.

Another objective was to have a model that was able to simulate the various yield curve dynamics encountered in practice, e.g. steepening, flattening and inversion. Figures 9 and 10 show historical yields up to 2002 followed by simulated yields for two years. Figure 9 demonstrates the model can simulate steepening and flattening, while Figure 10 demonstrates that the model can simulate inversion.

5. Conclusions

The objective of this paper is to identify a model that captures the salient features of the whole term structure, rather than one that just focuses on the short-term interest rate. It also has to be tractable in order to form a basis for asset pricing applications and forward simulations. To this end, we consider a three-factor continuous-time model within the affine class with a closed-form solution for the bond prices.

For our empirical analysis, the model is expressed in a state-space formulation which allows us to take into account both the cross-sectional and time-series information contained in the term structure data and we use the Kalman filter to estimate the parameters.

The model explains the cross-section of interest rates well with reasonably small yield errors. We also show that in forward simulations this model gives rise to a wide and realistic range of future interest rate scenarios, as shown by both the backtest and the simulations involving flattening / steepening / inversion of the yield curve.

Acknowledgements

The authors would like to express their gratitude to Pioneer Investments for their continuing support. In particular we would like to express our appreciation and thanks to Francesco Sandrini and Mark Scrowston. We are grateful to Professor Dempster for his many helpful suggestions, which greatly improved the final version of this paper.

References

Andersen, T.G., Benzoni, L., Lund, J., 2004. Stochastic volatility, mean drift, and jumps in the short-term interest rate. Working Paper, Carlson School of Management, University of Minnesota, Minneapolis.

Babbs, S.H., Nowman, K.B., 1999. Kalman filtering of generalized Vasicek term structure models. Journal of Financial and Quantitative Analysis 34 (1), 115--130.

Cairns, A.J.G., 2004. A family of term-structure models for long-term risk management and derivative pricing. Mathematical Finance 14 (3), 415--444.

Chan, K.C., Karolyi, G.A., Longstaff, F.A., Sanders, A.B., 1992. An empirical comparison of alternative models of the short-term interest rate. Journal of Finance 47 (3), 1209--1227.

Chen, R., Scott, L., 1993. ML estimation for a multifactor equilibrium model of the term structure. Journal of Fixed Income 3, 14--31.

Cox, J.C., Ingersoll, J.E., Ross, S.A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385--407.

Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. Journal of Finance 55 (5), 1943--1978.

De Jong, F., 2000. Time-series and cross-section information in affine term- structure models. Journal of Business and Economic Statistics 18 (3), 300--314.

De Jong, F., Santa-Clara, P., 1999. The dynamics of the forward interest rate curve: a formulation with state variables. Journal of Financial and Quantitative Analysis 34 (1), 131--157.

Dempster, A.P., Larid, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM-algorithm. Journal of the Royal Statistical Society: Series B 39, 1--38.

Duffie, D., Kan, R., 1996. A yield-factor model of interest rates. Mathematical Finance 6 (4), 379--406.

Geyer, A.L.J., Pichler, S., 1999. A state-space approach to estimate and test Cox-Ingersoll-Ross models of the term structure. Journal of Financial Research 22 (1), 107--130.

Harvey, A.C., 1993. Time Series Models. Hemel Hempstead: Harvester Wheatsheaf.

Heath, D., Jarrow, R., Morton, A., 1992. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica 60 (1), 77--105.

Ho, T.S.Y., Lee, S., 1986. Term structure movements and pricing interest rate contingent claims. Journal of Finance 41 (5), 1011--1029.

Hull, J.C., White, A.D., 1990. Pricing interest rate derivative securities. Review of Financial Studies 3 (4), 573--592.

James, J., Webber, N., 2000. Interest Rate Modelling. Wiley, Chichester.

Jamshidian, F., 1997. LIBOR and swap market models and measures. Finance & Stochastics 1 (4), 293--330.

Jones, D.R., Perttunen, C.D., Stuckmann, B.E., 1993. Lipschitzian optimization without the Lipschitz constant. Journal of Optimization Theory and Application 79 (1), 157--181.

Langetieg, T.C., 1980. A multivariate model of the term structure. Journal of Finance 35 (1), 71--97.

Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. Journal of Fixed Income 1, 54--61.

Litterman, R., Scheinkman, J., Weiss, L., 1991. Volatility and the yield curve. Journal of Fixed Income 1, 49--53.

Powell, M.J.D., 1964. An efficient method of finding the minimum of a function of several variables without calculating derivatives. Computer Journal 11, 302--304.

Rebonato, R., Mahal, S., Joshi, M., Buchholz, L.-D., Nyholm, K., 2005. Evolving yield curves in the real-world measures: a semi-parametric approach. Journal of Risk 7 (3), 29--61.

Vasicek, O., 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5 (2), 177--188.

Wilkie, A.D., Waters, H.R., Yang, S., 2004. Reserving, pricing and hedging for policies with guaranteed annuity options. British Actuarial Journal 10 (1), 101--152.

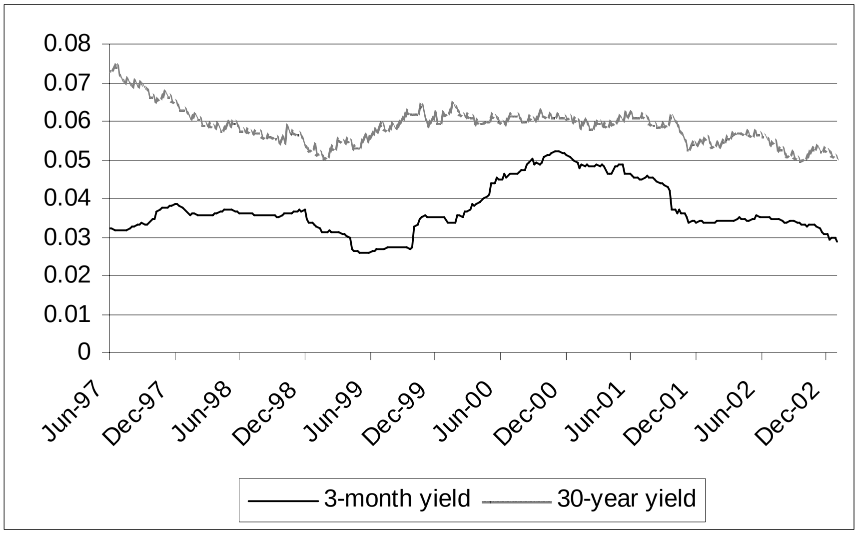

Figure 1: 3-Month and 30-Year EU Yields for the Period June 1997 – Dec 2002

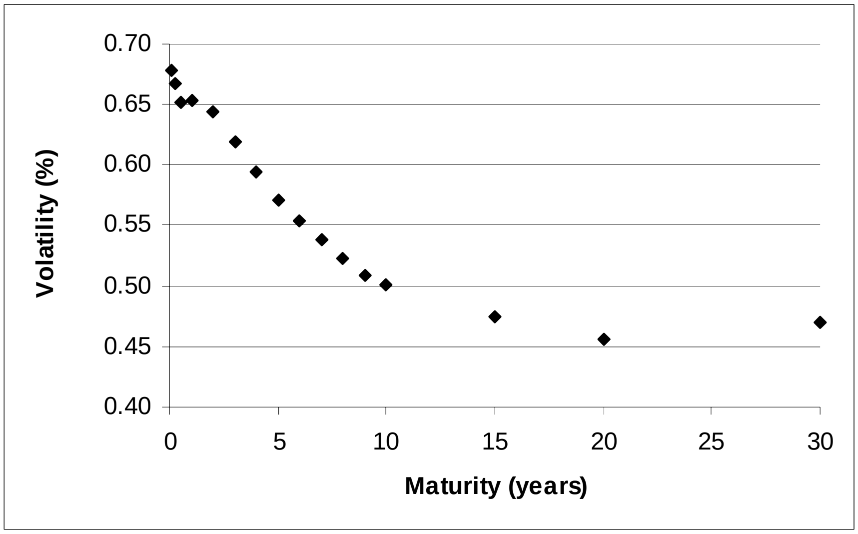

Figure 2: Weekly Standard Deviation of Yields for the period June 1997 – Dec 2002

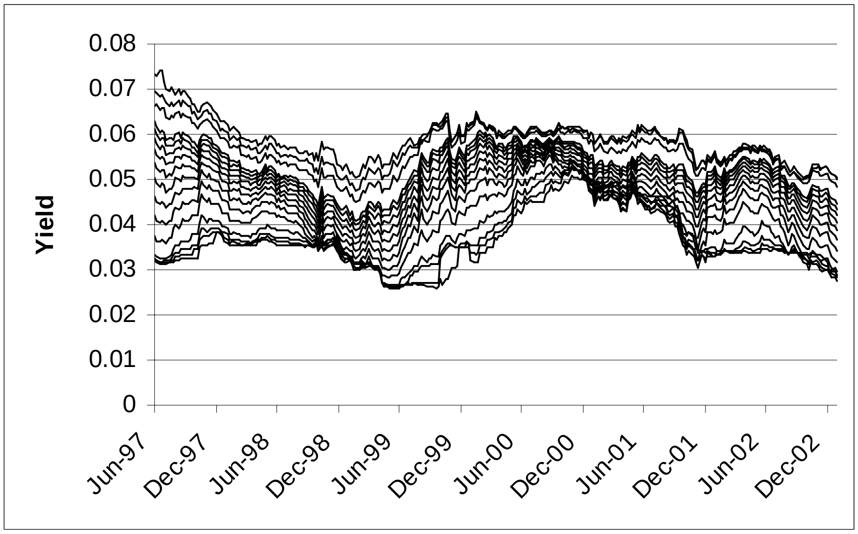

Figure 3: Historical Swap Yields of Varying Maturities over the Period

June 1997- December 2002

Euro Data		Estimated Value	Standard Error
Long term risk neutral mean *X*	μ_X/λ_X	0.199	1.69E-04
Long term risk neutral mean *Y*	μ_Y/λ_Y	-0.134	1.69E-04
Speed of mean reversion *X*	λ_X	0.161	1.03E-03
Speed of mean reversion *Y*	λ_Y	1.332	6.87E-03
Speed of mean reversion *R*	k	0.117	1.64E-03
Volatility *X*	σ_X	0.030	1.89E-04
Volatility *Y*	σ_Y	0.186	9.80E-04
Volatility *R*	σ_R	0.006	2.26E-04
Correlation *X* and *Y*	ρ_{XY}	-0.642	6.94E-03
Correlation *X* and *R*	ρ_{XR}	0.177	1.82E-02
Correlation *Y* and *R*	ρ_{YR}	-0.540	1.81E-02
Market price of risk for *X*	γ_X	0.556	3.91E-03
Market price of risk for *Y*	γ_Y	-1.017	5.50E-03
Market price of risk for *R*	γ_R	0.096	1.65E-02

Table 1: Estimated Parameters Using Kalman Filter

	Maturity	Estimated Value	Standard Error
$\sqrt{h_1}$	1 month	1.57E-03	6.63E-05
$\sqrt{h_2}$	3 months	8.64E-04	3.81E-05
$\sqrt{h_3}$	6 months	1.55E-04	3.19E-05
$\sqrt{h_4}$	1 year	6.71E-04	2.96E-05
$\sqrt{h_5}$	2 years	5.08E-04	2.15E-05
$\sqrt{h_6}$	3 years	2.85E-04	1.21E-05
$\sqrt{h_7}$	4 years	1.49E-04	7.03E-06
$\sqrt{h_8}$	5 years	4.96E-05	4.59E-06
$\sqrt{h_9}$	6 years	6.58E-05	2.89E-06
$\sqrt{h_{10}}$	7 years	1.00E-05	3.83E-06
$\sqrt{h_{11}}$	8 years	9.44E-05	4.1E-06
$\sqrt{h_{12}}$	9 years	1.75E-04	7.63E-06
$\sqrt{h_{13}}$	10 years	2.94E-04	1.28E-05
$\sqrt{h_{14}}$	15 years	7.45E-04	3.14E-05
$\sqrt{h_{15}}$	20 years	1.23E-03	5.32E-05
$\sqrt{h_{16}}$	30 years	2.37E-03	1.03E-04

Table 2: Measurement Errors

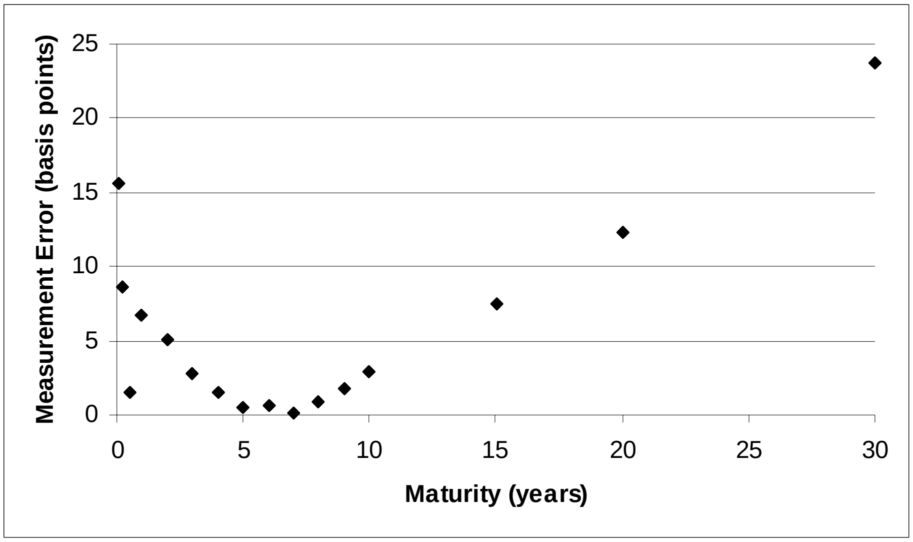

Figure 4: Measurement Error of the Fitted Yields

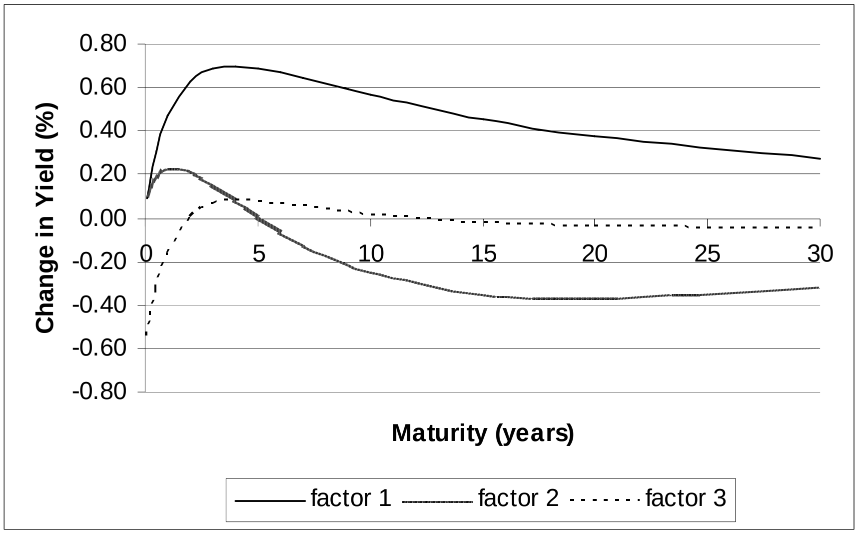

Figure 5: Factor Loading of the Three-Factor Model

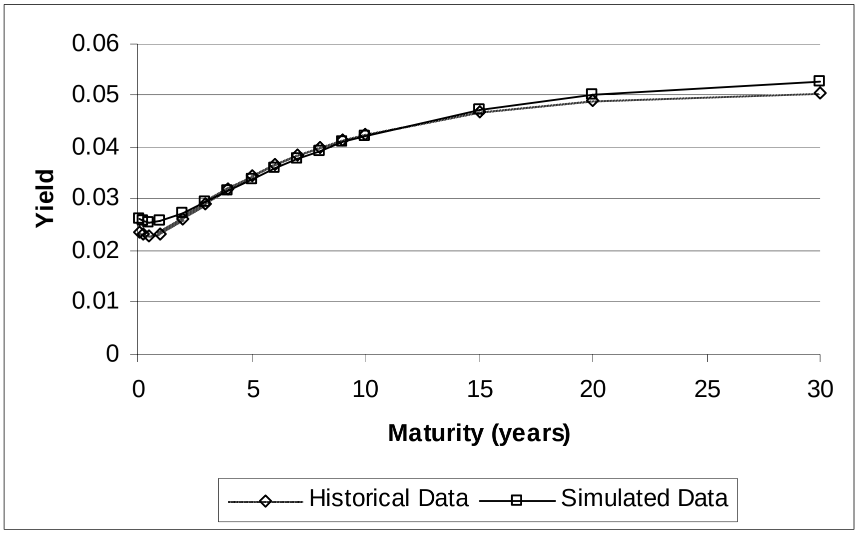

Figure 6: Mean Level of Yields over 2003 for Historical and Simulated Data

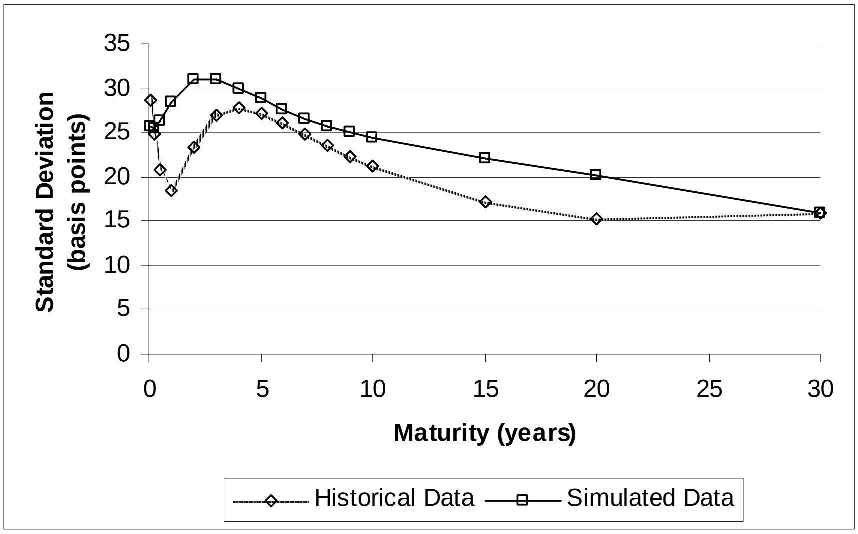

Figure 7: Standard Deviation of Yields over 2003 for Historical and Simulated Data

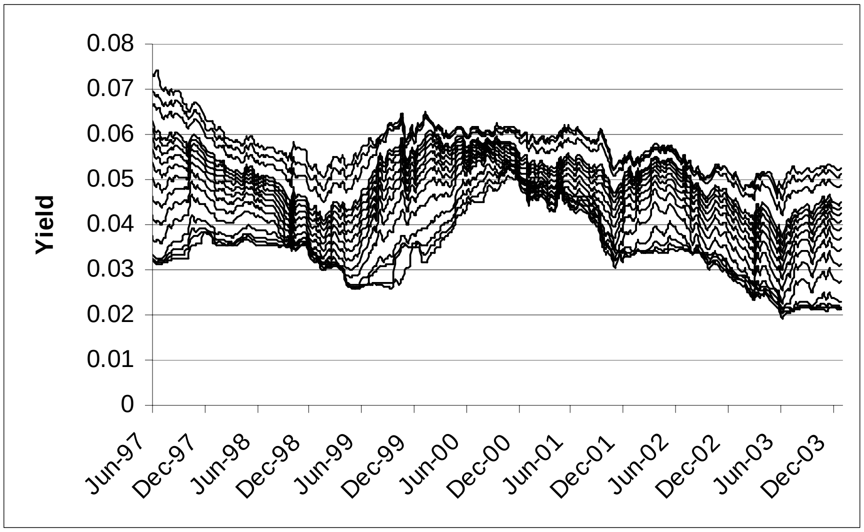

Figure 8: Historical Yields for June 1997 – Dec 2003

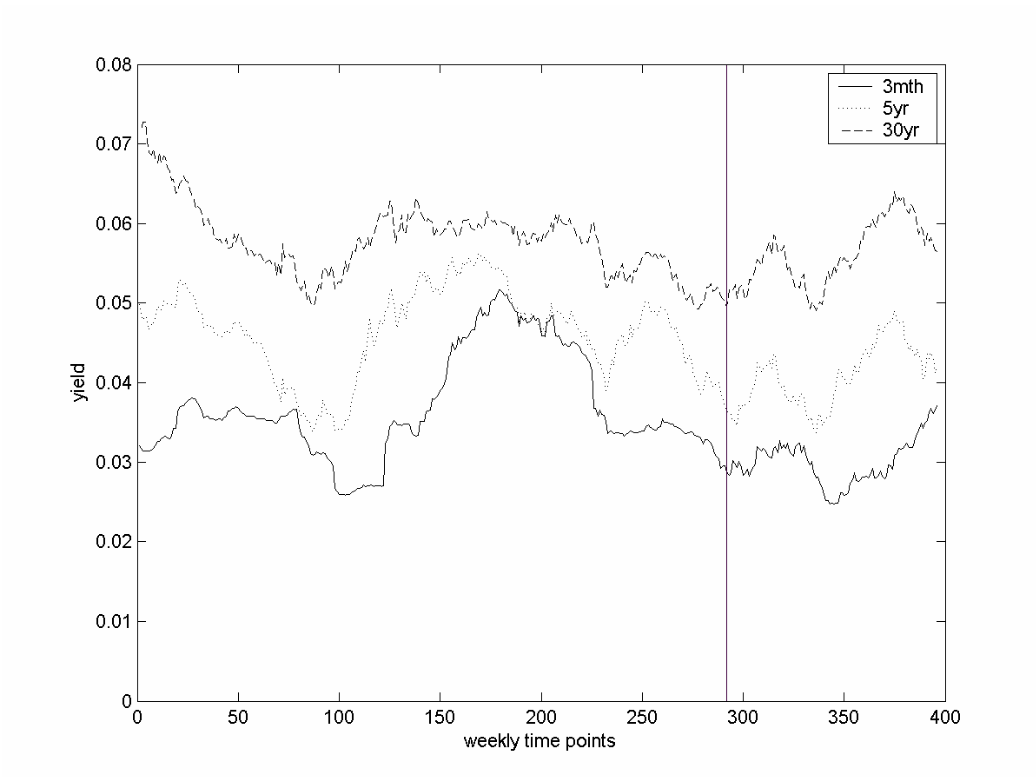

Figure 9: Forward Simulation Showing Steepening and Flattening of the Yield Curve

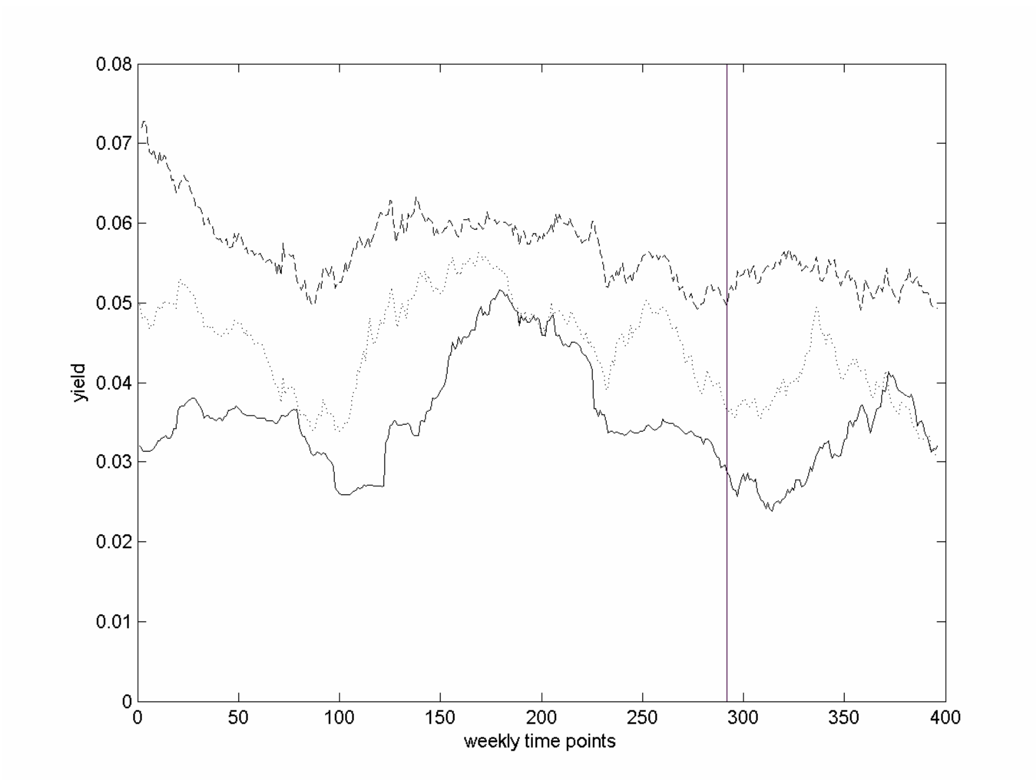

Figure 10: Forward Simulation Showing Inversion of the Yield Curve