QUANTITATIVE FINANCE
RESEARCH CENTRE

UNIVERSITY OF
TECHNOLOGY SYDNEY

QUANTITATIVE FINANCE
RESEARCH CENTRE

UTS

THINK.CHANGE.DO

QUANTITATIVE FINANCE RESEARCH CENTRE

Research Paper 151 January 2005

The Volatility Structure of the Fixed Income Market under the HJM Framework: A Nonlinear Filtering Approach

Carl Chiarella, Hing Hung and Thuy-Duong To

ISSN 1441-8010 www.qfrc.uts.edu.au

THE VOLATILITY STRUCTURE OF THE FIXED INCOME MARKET UNDER THE HJM FRAMEWORK: A NONLINEAR FILTERING APPROACH

CARL CHIARELLA+, HING HUNG‡, AND THUY-DUONG TÔ*

School of Finance and Economics
University of Technology, Sydney
1-54 Quay Street
Broadway NSW 2007
Australia
+carl.chiarella@uts.edu.au
‡hing.hung@uts.edu.au
*td.to@uts.edu.au

ABSTRACT. This paper considers the dynamics for interest rate processes within a multi-factor Heath, Jarrow and Morton (1992) specification. Despite the flexibility of and the notable advances in theoretical research about the HJM models, the number of empirical studies is still inadequate. This paucity is principally because of the difficulties in estimating models in this class, which are not only high-dimensional, but also nonlinear and involve latent state variables. This paper treats the estimation of a fairly broad class of HJM models as a nonlinear filtering problem, and adopts the local linearization filter of Jimenez and Ozaki (2003), which is known to have some desirable statistical and numerical features, to estimate the model via the maximum likelihood method. The estimator is then applied to the interbank offered-rates of the U.S, U.K, Australian and Japanese markets. The two-factor model, with the factors being the level and the slope effect, is found to be a reasonable choice for all of the markets. However, the contribution of each factor towards overall variability of the interest rates and the financial reward each factor claims differ considerably from one market to another.

Key words: Term structure; Heath-Jarrow-Morton; Local Linearization; Filtering;

JEL classifications: C51, E43, G12

Date: Version: Jan 2005.
* Corresponding author.

1. Introduction

Management of interest rate risk is of crucial importance to financial institutions and corporations. The volatility structure of this interest rate market plays a crucial role in assessing and managing the value as well as the risk of bond and interest rate derivative portfolios. Various interest rate models have been considered, amongst which the Heath-Jarrow-Morton (1992) (hereafter HJM) framework provides a very flexible framework for interest rate modelling. Despite its nice theoretical flexibility, the application of the HJM class of models to practical problems is hindered by the difficulty of model estimation. This is principally due to the fact that the underlying state variables of the HJM models are un-observable quantities, and the dynamics are usually non-Markovian and non-linear in their (latent) state variables.

Theoretical research on HJM models has shown that for a fairly broad family of volatility functions, the underlying stochastic system can be Markovianized, and thereby easing the computational complexity involved. However, the problems of nonlinearity and the existence of latent variables still exist, and the empirical analysis of HJM models has centered around certain volatility functions that lead to convenient properties for the system, for example, the class of affine or square root affine volatilities.

It should also be noted that the estimation for stochastic models is already a challenging task for systems with affine or square root affine volatilities. Duffee and Stanton (2004) analyze the performance of different estimation methods for dynamic term structure models. They find that the standard maximum likelihood estimator (MLE) does a very poor job of estimating the parameters that determine expected changes in interest rates. Furthermore they find that the efficient method of moment (EMM) estimator is an unacceptable alternative, even where the MLE performs well. They conclude that the Kalman filter is a reasonable choice, even in the non-Gaussian setting where the filter is not exact. In that case, they advocate the use of a variant of the Kalman filter, where the updating equation for the state variables is a linearized version of the drift using its first derivative.

In light of the findings of Duffee and Stanton (2004) this paper pursues further the filtering approach. To deal with the nonlinear nature of the problem we advocate the use of the local linearization filter of Jimenez and Ozaki (2002, 2003). The main idea is to linearize the system dynamics according to the Itô formula, utilizing both the drift and the diffusion terms, to better take into account the stochastic behaviour of the system, and then to apply the (readily available) optimal linear filter. We have chosen this filter as it has been shown by Shoji (1998) to have good bias properties and by Jimenez et al. (1999) to have a number of computational advantages. The estimation method is able to exploit both the time series and cross sectional information of the yield curve.

We propose and motivate a three-factor volatility specification and apply the local linearization filter to analyze the volatility structure of the interbank offered-rates in the U.S., the U.K, the Australian and the Japanese markets. These markets have been chosen to represent different regions in the world. The rest of the paper is organized as follows. Section 2 introduces the model. The econometric implication of the model and the proposed estimation method are discussed in Section 3. Empirical results are then presented in Section 4, and Section 5 concludes the paper.

2. Model Framework

The general framework for the interest rate models considered in this paper is introduced in Heath, Jarrow and Morton (1992), where the instantaneous forward rates $r(t,x)$ (the rate that can be contracted at time t for instantaneous borrowing/lending at future time $t+x$) are assumed to satisfy SDEs of the form[1]

$$\begin{aligned} r(t,x) = r(0,t+x) &+ \int_0^t \boldsymbol{\sigma}(s,t+x)' \left[\bar{\boldsymbol{\sigma}}(s,t+x) - \boldsymbol{\phi}(s)\right] ds \\ &+ \int_0^t \boldsymbol{\sigma}(s,t+x)' d\boldsymbol{W}(s), \end{aligned} \tag{2.1}$$

where

$$\bar{\boldsymbol{\sigma}}(s,t+x) = \int_s^{t+x} \boldsymbol{\sigma}(s,u)du,$$

and $\boldsymbol{\sigma}(t,x)$, $\boldsymbol{\phi}(t)$ are I-dimensional processes and $\boldsymbol{W}(t)$ is a standard I-dimensional vector of independent Wiener processes under the market measure $\mathcal{P}$, $I \in \mathbb{N}_+$ and the superscript $'$ represents matrix transposition. The vector $\boldsymbol{\phi}(t)$ can be interpreted as the market price of interest rate risk vector associated with $d\boldsymbol{W}(t)$. In general, $\boldsymbol{\sigma}$ and $\boldsymbol{\phi}$ may depend on a number of forward rates $r(t,x)$.[2]

The HJM model framework is chosen as it yields arbitrage-free models that fit the initial yield curve by construction. The subclass of HJM models which are particularly suited to practical implementation are those which can be Markovianized. Carverhill (1994), Ritchken and Sankarasubramanian (1995), Bhar and Chiarella (1997a), Inui and Kijima (1998), de Jong and Santa-Clara (1999) and Björk and Svensson (2001) discuss various specifications of the forward rate volatilities $\sigma(t,x)$ that lead to Markovian representations of the forward rate dynamics. Chiarella and Kwon

---

[1]We are in fact using the Brace et al. (1997) implementation of the HJM model. This is more appropriate to capture the dynamics of LIBOR and various other market quoted rates.

[2]In this notation, $r(t,0)$ denotes the instantaneous rate of interest that we henceforth write as $r(t)$.

(2001b, 2003) introduce a specification that leads to a fairly broad and convenient class of models. The models in this class satisfy the assumption:

Assumption 2.1. *(i) For each $1 \leq i \leq I$, there exists $L_i \in \mathbb{N}$ such that the components, $\sigma_i(t,x)$, of the forward rate volatility process have the form*

$$\sigma_i(t,x) = \sum_{l=1}^{L_i} c_{il}(t)\sigma_{il}(x) \tag{2.2}$$

where $c_{il}(t)$ are stochastic processes and $\sigma_{ij}(x)$ are deterministic functions.

(ii) There exist $M \in \mathbb{N}$ and a sequence $x_1 < \cdots < x_M \in \mathbb{R}_+$ such that the processes $c_{il}(t)$ have the form

$$c_{il}(t) = \hat{c}_{il}(t, r(t,x_1), \ldots, r(t,x_M)), \tag{2.3}$$

where $\hat{c}$ is deterministic in its arguments.

Chiarella and Kwon (2003) then prove that the forward curve can be expressed as an affine function of a set of N discrete tenor forward rates

$$\boldsymbol{r}(t,\tau_1,\ldots,\tau_N) = [r(t,\tau_1),\ldots,r(t,\tau_N)]'$$

(see Appendix A for a brief summary). This set of forward rates forms a Markov process. In terms of the real world measure, where $\boldsymbol{\phi} \equiv (\phi_1,\ldots,\phi_I)$ is the vector of market prices of risk associated with the Wiener process $\boldsymbol{W}$, the system of stochastic differential equations for the instantaneous forward rates becomes[3]

$$\begin{aligned} dr(t,x) =& [p_0(t,x,\tau_1,\ldots,\tau_N) + \boldsymbol{p}_1'(t,x,\tau_1,\ldots,\tau_N)\boldsymbol{r}(t,\tau_1,\ldots,\tau_N) \\ & - \boldsymbol{\phi}'\boldsymbol{\sigma}(t,t+x)]dt + \boldsymbol{\sigma}(t,t+x)'d\boldsymbol{W}(t). \end{aligned} \tag{2.4}$$

The yield $y(t,x)$ on the $(t+x)$-maturity zero coupon bond can be calculated from the instantaneous forward rates via

$$y(t,x) = \frac{1}{x}\int_0^x r(t,u)du, \tag{2.5}$$

and can also be expressed as an affine function of the forward rates, that we write in the form

$$y(t,x) = q_0(t,x,\tau_1,\ldots,\tau_N) - \boldsymbol{q}'(t,x,\tau_1,\ldots,\tau_N)\boldsymbol{r}(t,\tau_1,\ldots,\tau_N), \tag{2.6}$$

where the $q_i(t,x,\tau_1,\ldots,\tau_N)$ is a set of deterministic functions[4]. We therefore have an affine term structure model. This model is not nested inside the popular affine model class considered in Duffie and Kan (1996), even though there will be occasions when the two classes overlap.

[3]For definition of the coefficient functions p_0 and $\boldsymbol{p}$, see Appendix A.

[4]Again see Appendix A for definitions of the q_i.

3. ESTIMATION FRAMEWORK

3.1. **The model specification.**

The empirical work of Litterman and Scheinkman (1991), Chen and Scott (1993), Knez et al. (1994), Singh (1995), who use principal component analysis, suggests that there are at most three factors affecting the volatility of interest rates. Guided by this insights we propose to use a three-dimensional Wiener process in the specification (2.1), with the corresponding volatility functions

$$\sigma_1(t,x) = \gamma_1 r^{\lambda}(t), \tag{3.1}$$

$$\sigma_2(t,x) = \gamma_2(r(t,\tau) - r(t)), \tag{3.2}$$

$$\sigma_3(t,x) = \gamma_3 \, \mathrm{e}^{-\kappa(x-t)} \, . \tag{3.3}$$

The first volatility function $\sigma_1(t,x)$ represents the level factor. If $\lambda_1 = 0.5$ we would obtain a Cox-Ingersoll-Ross (1985a) type of volatility. The second volatility function $\sigma_2(t,x)$ reflects the influence of the slope of the yield curve on interest rate volatility, with the difference $r(t,\tau)-r(t)$ proxying the slope. Finally, the last volatility function $\sigma_3(t,x)$ allows a shock in the corresponding Wiener process to have different impacts at different maturities along the yield curve.

The market price of risk terms ϕ_1, ϕ_2, ϕ_3 are assumed to follow a square-root type of processes, ie. they are mean reverting and have volatility functions proportional to the square root of their own levels, ie.

$$d\phi_i = \alpha_i(\bar{\phi}_i - \phi_i)dt + \beta_i\sqrt{\phi_i(t)}dW_i(t). \tag{3.4}$$

Intuitively, the specification suggests that the market prices of different interest rate risks are always positive and tend to converge to their long run equilibria.

3.2. **Econometric implication of the model.**

Some similar and other specialized models of the HJM class considered here have been empirically analyzed. Bliss and Ritchken (1996) consider the case where the volatility function in (2.2) can be written as[5]

$$\sigma(t,x) = c(t) \, \mathrm{e}^{-\kappa x} \, .$$

This specification covers our single-factor model, as each of our volatility function can be written in that form. For example, with $\sigma_1(t,x) = \gamma_1 r^{\lambda}(t)$, the value of κ is zero and $c(t) = \gamma_1 r^{\lambda}(t)$. The key idea of their approach is to exploit the relationship (2.6) for the yields, into which they introduce an error term, then estimate their model via

[5]With this volatility function, the model can be Markovianized using two state variables.

the Maximum Likelihood procedure [6]. The main drawback of this approach is that the estimation procedure can only identify κ, as the relationship (2.6) does not depend on the parameters characterizing function $c(t)$ (γ_1 and λ in our example). However, all of the parameters in the models are important in practical work, such as the determination of the price of a derivative contract.

de Jong and Santa-Clara (1999) also empirically study two-state variable HJM models where the volatility function of the system is proportional to the square root of the state variables. However, they overcome the disadvantages of Bliss and Ritchken (1996) approach by using both the relationship (2.6) and the Markovian system (2.4) in their estimation procedure. They use the Kalman filtering method where (2.6) serves as the observation equation and (2.4) is discretized into a state transition equation. In a more general setting, it is not clear how to discretize the structural stochastic system, and the behaviour of the estimator is clearly dependent on the method used in this discretization.

In this paper, we advocate the local linearization filter (hereafter the LL filter) of Jimenez and Ozaki (2002, 2003). This approach is still based on the Kalman filter for a discrete linear system. However, Jimenez and Ozaki do not discretize the nonlinear system directly, but rather approximate it by a system linear in both its drift and its diffusion terms, for which a linear Kalman filter turns out to be readily applicable. The approximation is not based on the first order Taylor approximation used in the standard extended Kalman filter framework, but is instead based on a second order approximation using the Itô formula to better take into account the stochastic behaviour of the underlying state variables.

In his comparative study, Shoji (1998) analyzed the performance of the maximum likelihood estimator based on the LL filter and the one based on the extended Kalman filter for a system with additive noise (i.e. the volatility function is not dependent on the state variables). Shoji used Monte Carlo simulation to show that the LL filter provided estimates with smaller bias, particularly in estimation of the coefficient of the drift term. Jimenez et al. (1999) compared the LL scheme with other linearization schemes for systems with either additive or multiplicative noise (i.e. the volatility function is dependent on the state variables). They also reported a number of numerical advantages of the LL filter, including numerical stability, better accuracy and the order of strong convergence.

[6] The relationship Bliss and Ritchken use is actually an expression of the whole yield curve as an affine function of some particular yields rather than the forward rates. This can be derived very simply from the model here.

3.3. **The local linearization filter and the maximum likelihood estimator.**

Consider the state space model defined by the continuous state equation

$$d\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t))dt + \sum_{i=1}^{m} \mathbf{g}_i(t, \mathbf{x}(t))dW_i(t), \tag{3.5}$$

and the discrete observation equation[7]

$$\mathbf{z}_{t_j} = \mathbf{C}(t_j)\mathbf{x}(t_j) + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J, \tag{3.6}$$

where $\mathbf{f}$ and $\mathbf{g}_i$ are nonlinear functions, $\mathbf{x}(t) \in \mathbb{R}^d$ is the state vector at the instant of time t, $\mathbf{z}_{t_j} \in \mathbb{R}^r$ is the observation vector at the instant of time t_j, $\boldsymbol{W}$ is an m-dimensional Wiener process, and $\{\mathbf{e}_{t_j} : \mathbf{e}_{t_j} \sim \mathcal{N}(0, \mathbf{\Pi}), j = 0, \ldots, J\}$ is a sequence of i.i.d. random vectors.

The system functions $\mathbf{f}$ and $\mathbf{g}_i$ can be linearly approximated. Jimenez and Ozaki (2003) proposed to approximate them via a truncated Ito-Taylor expansion to better take into account the stochastic behaviour of the underlying state system. For example, the approximation for $\mathbf{f}$ is

$$\begin{aligned}\mathbf{f}(t, \mathbf{x}(t)) \approx & f(s, \mathbf{u}) + \left(\frac{\partial \mathbf{f}(s, \mathbf{u})}{\partial s} + \frac{1}{2} \sum_{k,l=1}^{d} [\mathbf{G}(s, \mathbf{u})\mathbf{G}'(s, \mathbf{u})]^{k,l} \frac{\partial^2 \mathbf{f}(s, \mathbf{u})}{\partial \mathbf{u}^k \partial \mathbf{u}^l} \right) (t - s) \\ & + \mathbf{J}_{\mathbf{f}}(s, \mathbf{u})(\mathbf{x}(t) - \mathbf{u}),\end{aligned} \tag{3.7}$$

where $(s, \mathbf{u}) \in \mathbb{R} \times \mathbb{R}^d$, $\mathbf{J}_{\mathbf{f}}(s, \mathbf{u})$ is the Jacobian of $\mathbf{f}$ evaluated at the point $(s, \mathbf{u})$ and $\mathbf{G}(s, \mathbf{u})$ is the $d \times m$ matrix defined by $\mathbf{G}(s, \mathbf{u}) \equiv (\mathbf{g}_1, \ldots, \mathbf{g}_m)$.

Using such approximations for $\mathbf{f}$ and $\mathbf{g}_i$, the solution of the nonlinear state equation (3.5) can be approximated by the solution of the piecewise linear stochastic differential

---

[7] A full (nonlinear) specification of the observation equation would be

$$\mathbf{z}_{t_j} = \mathbf{h}(t_j, \mathbf{x}(t_j)) + \sum_{i=1}^{n} \mathbf{p}_i(t_j, \mathbf{x}(t_j))\xi^i_{t_j} + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J,$$

where $\mathbf{h}$ and $\mathbf{p}_i$ are nonlniear functions, $\{\boldsymbol{\xi}_{t_j} : \boldsymbol{\xi}_{t_j} \sim \mathcal{N}(0, \mathbf{\Lambda}), \mathbf{\Lambda} = \mathrm{diag}((\lambda_1, \ldots, \lambda_n)), j = 0, \ldots, J\}$ is a sequence of random vector i.i.d., and $\xi^i_{t_j}$ and $\mathbf{e}_{t_j}$ are uncorrelated for all i and j. However, in most finance applications, including ours, a linear specification for $\mathbf{h}$ is all that is required and there is no need to include the extra noise term $\boldsymbol{\xi}$.

equation[8]

$$d\mathbf{y}(t) = \Big(\mathbf{A}(t_j, \hat{\mathbf{y}}_{t_j|t_j})\mathbf{y}(t) + \mathbf{a}(t, t_j, \hat{\mathbf{y}}_{t_j|t_j})\Big)\, dt + \sum_{i=1}^{m} \Big(\mathbf{B}_i(t_j, \hat{\mathbf{y}}_{t_j|t_j})\mathbf{y}(t) + \mathbf{b}_i(t, t_j, \hat{\mathbf{y}}_{t_j|t_j})\Big)\, dW_i(t) \tag{3.8}$$

for all $t \in [t_j, t_{j+1})$, starting at $\mathbf{y}(t_0) = \hat{\mathbf{y}}_{t_0|t_0} = \hat{\mathbf{x}}_{t_0|t_0}$. The various quantities appearing in (3.8) are defined as

$$\begin{aligned}
\hat{\mathbf{x}}_{t|\rho} &= \mathbb{E}\left(\mathbf{x}(t)|Z_\rho\right), \quad Z_\rho = \{\mathbf{z}_{t_j} : t_j \le \rho\},\\
\hat{\mathbf{y}}_{t|\rho} &= \mathbb{E}\left(\mathbf{y}(t)|Z_\rho\right),\\
\mathbf{A}(s, \mathbf{u}) &= \mathbf{J}_{\mathbf{f}}(s, \mathbf{u}),\\
\mathbf{B}_i(s, \mathbf{u}) &= \mathbf{J}_{\mathbf{g_i}}(s, \mathbf{u}),\\
\mathbf{a}(t, s, \mathbf{u}) &= \mathbf{f}(s, \mathbf{u}) - \mathbf{J}_{\mathbf{f}}(s, \mathbf{u})\mathbf{u} + \frac{\partial \mathbf{f}(s, \mathbf{u})}{\partial s}(t - s)\\
&\quad + \frac{1}{2}\sum_{k,l=1}^{d} [\mathbf{G}(s, \mathbf{u})\mathbf{G}'(s, \mathbf{u})]^{k,l} \frac{\partial^2 \mathbf{f}(s, \mathbf{u})}{\partial \mathbf{u}^k \partial \mathbf{u}^l}(t - s),\\
\mathbf{b}_i(t, s, \mathbf{u}) &= \mathbf{g}_i(s, \mathbf{u}) - \mathbf{J}_{\mathbf{g}_i}(s, \mathbf{u})\mathbf{u} + \frac{\partial \mathbf{g}_i(s, \mathbf{u})}{\partial s}(t - s)\\
&\quad + \frac{1}{2}\sum_{k,l=1}^{d} [\mathbf{G}(s, \mathbf{u})\mathbf{G}'(s, \mathbf{u})]^{k,l} \frac{\partial^2 \mathbf{g}_i(s, \mathbf{u})}{\partial \mathbf{u}^k \partial \mathbf{u}^l}(t - s).
\end{aligned}$$

The approximate stochastic differential equation (3.8) and the corresponding observation equation (see (3.6))

$$\mathbf{z}_{t_j} = \mathbf{C}(t_j)\mathbf{y}(t_j) + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J, \tag{3.9}$$

form a linear state space system. The optimal linear filter proposed by Jimenez and Ozaki (2002) can be applied (see Appendix B for its definition) to determine the conditional mean $\hat{\mathbf{y}}_{t|\rho}$ and conditional covariance matrix $\mathbf{P}_{t|\rho} = \mathbb{E}((\mathbf{y}(t) - \hat{\mathbf{y}}_{t|\rho})(\mathbf{y}(t) - \hat{\mathbf{y}}_{t|\rho})'|Z_\rho)$ for all $\rho \le t$.

Due to the assumption of multivariate normality of the disturbances $\mathbf{e}_{t_j}$ (and if the initial state vector also has a proper multivariate normal distribution), the distribution of $\mathbf{z}_{t_{j+1}}$ conditional on Z_{t_j} is itself normal (see (3.9)). The mean and covariance matrix of this conditional distribution are given directly by the local linearization filter above. Therefore, a maximum likelihood estimator for the model parameters can be easily derived.

[8] We use $\mathbf{y}(t)$ to denote the solution to the approximate system to distinguish it from $\mathbf{x}(t)$ the solution to the true system.

Let $\boldsymbol{\theta}$ be the vector of parameters of interest, which include all parameters specifying the state space model (3.8) and (3.9), plus the initial state values of $\hat{\mathbf{x}}_{t_0|t_0}$ and $\mathbf{P}_{t_0|t_0}$. The log likelihood function for $\mathbf{Z}$ is

$$\mathcal{L}_{\mathbf{Z}}(\boldsymbol{\theta}) = -\frac{rJ}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{J}\ln|\boldsymbol{\Sigma}_{t_j}| - \frac{1}{2}\sum_{j=1}^{J}\boldsymbol{\nu}'_{t_j}\boldsymbol{\Sigma}_{t_j}^{-1}\boldsymbol{\nu}_{t_j} \tag{3.10}$$

where the innovation equations are

$$\boldsymbol{\nu}_{t_j} = \mathbf{z}_{t_j} - \mathbf{C}(t_j)\hat{\mathbf{y}}_{t_j|t_{j-1}}, \tag{3.11}$$

$$\boldsymbol{\Sigma}_{t_j} = \mathbf{C}(t_j)\mathbf{P}_{t_j|t_{j-1}}\mathbf{C}'(t_j) + \boldsymbol{\Pi}. \tag{3.12}$$

The maximum likelihood estimator of $\boldsymbol{\theta}$ is then

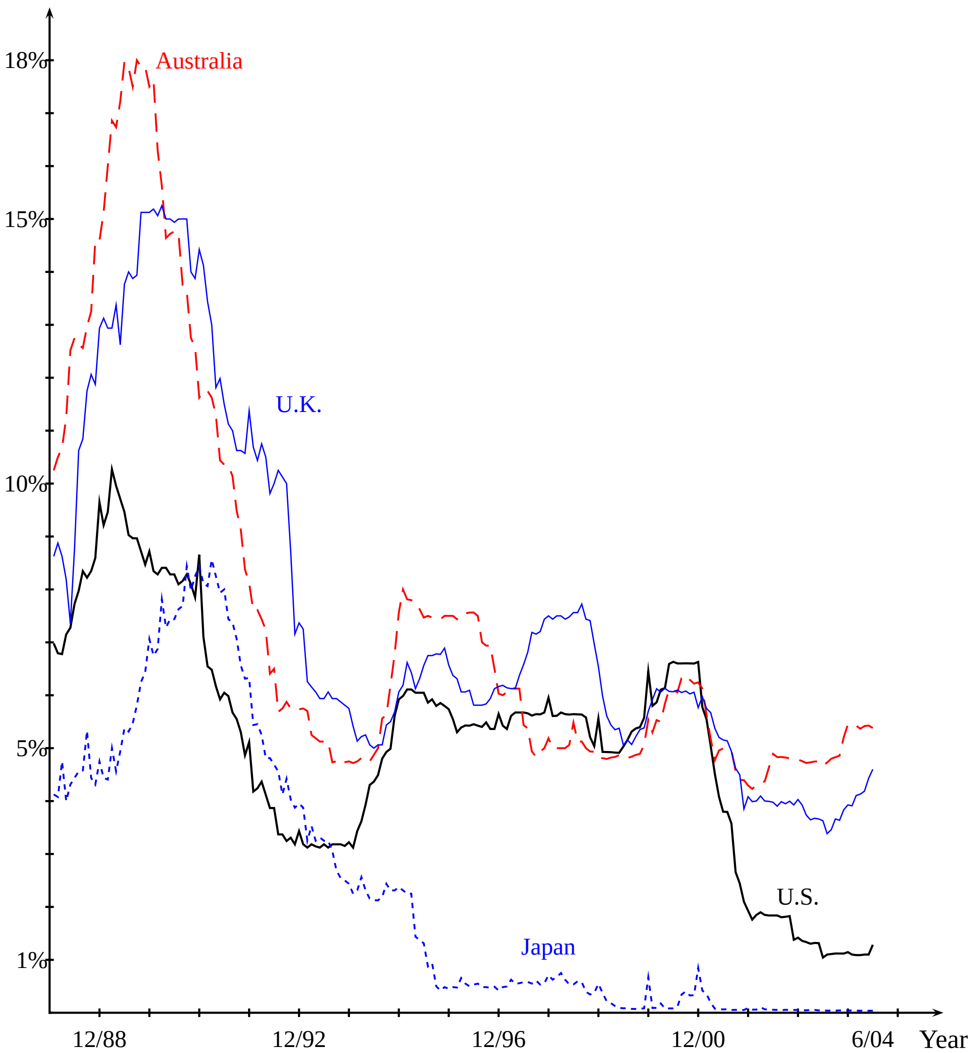
$$\hat{\boldsymbol{\theta}} = \max_{\boldsymbol{\theta}} \mathcal{L}_{\mathbf{Z}}(\boldsymbol{\theta}). \tag{3.13}$$

3.4. **Econometric implementation.**

We now view our model as a continuous-discrete nonlinear state space system, where (2.4) and (3.4) serve as the nonlinear state equations, and (2.6) serves as the linear (affine) observation equation. Similar to the standard practice in the literature, we introduce into the observation equation a measurement error, which reflects the fact that the model cannot fit all observed yields simultaneously. This measurement error is assumed to follow a multivariate normal distribution. The local linearization filter can be readily applied to yield the maximum likelihood estimator of $\boldsymbol{\theta}$, the vector of parameters of interest, which includes all of the parameters of the volatility functions (3.1) - (3.3), of the market price of risk specification (3.4) and the initial conditional mean vector $\hat{\mathbf{x}}_{t_0|t_0}$ and conditional variance matrix $\mathbf{P}_{t_0|t_0}$.

The numerical difficulties associated with any estimation procedures for stochastic systems are well-known. Amongst them, system stability, matrix inversion to calculate the likelihood function, convergence of the optimization routine and significance of the estimates are the main problems. To partly overcome these problems, we maximize the likelihood function using a genetic algorithm (Holland (1975), Mitchell (1996), Vose (1999), Michalewicz (1999)). Genetic algorithms use the evolutionary principle to solve difficult problems with objective functions that do not possess “nice” properties such as continuity and differentiability. The algorithms search the solution space of a function, and implement a “survival of the fittest” strategy to improve the solutions.

FIGURE 1. 1-month interbank offered-rates



4. EMPIRICAL ANALYSIS

4.1. **The Data.**

We estimate the model using the interbank offered-rates in the U.S, U.K, Australian and Japanese markets downloaded from Datastream®. The data consists of monthly observations for contracts with maturity from 1 month to 12 months, spanning a period from January 1988 to June 2004.

Figure 1 shows the 1-month rates for different markets. Over the 16-year period, interest rates change significantly. The overall pattern is an increasing trend for the last years of the 80s, followed by a sharp decrease throughout the first half of the 90s.

In the second half of the 90s, interest rates moved considerably around a temporary "long term" level, before decreasing during the year 2001. The rates in Australia and U.K. then picked up again, whereas the U.S. still experienced a decline in rates, and the Japanese rates stayed at a very low level below 10 basis points. All of the rates display a high level of autocorrelation, as can be seen in Table 1.

TABLE 1. The 1-month interbank offered-rates

	U.S.	Australia	U.K.	Japan
Mean	5.09%	7.55%	7.54%	2.34%
Standard deviation	2.27 %	3.80%	3.37%	2.65%
AC(1)	0.9832	0.9938	0.9916	0.9922

We also analyzed the principal components of the zero yield curve constructed from the interbank offered-rates. In all of the markets, three components are able to explain 100% of the variation in the yields, however the last component plays a very negligible role, only explaining 0.01%-0.02% of the total variation, as reported in Table 2.

TABLE 2. Principal component analysis of zero yield curves

% variation explained	U.S.	Australia	U.K.	Japan
Principal component 1	99.64	99.76	99.68	99.87
Principal component 2	0.34	0.23	0.31	0.11
Principal component 3	0.02	0.01	0.01	0.01

4.2. **Empirical Results.**

To increase the computational accuracy we estimated the 1-factor, 2-factor and 3-factor models separately, as follows:

- 1-factor model

$$\sigma_1(t,x) = \gamma_1 r^{\lambda}(t)$$

- 2-factor model

$$\sigma_1(t,x) = \gamma_1 r^{\lambda}(t),$$
$$\sigma_2(t,x) = \gamma_2(r(t,\tau) - r(t))$$

- 3-factor model

$$\sigma_1(t,x) = \gamma_1 r^{\lambda}(t),$$
$$\sigma_2(t,x) = \gamma_2(r(t,\tau) - r(t)),$$
$$\sigma_3(t,x) = \gamma_3\,\mathrm{e}^{-\kappa(x-t)}\,.$$

We expected the third factor to contribute very slightly to the total variation of the yield curve, and therefore including it may make the task of separating different components harder. There are also a lot more parameters involved in the 3-factor model, which may cause difficulties in the numerical optimization of the likelihood function.

4.2.1. *The U.S market.*

The parameter estimates for the U.S. market can be found in Table 3. All of the estimates are highly significant. However, the numerical optimization routine fails to find a higher likelihood function for the 3-factor model.

For the 1-factor model (whose interest rate volatility function is $\gamma_1 r^\lambda$), the estimate of λ is 1.97, higher than the value close to 1.5 found by Chan et al. (1992), and the range of 0.5 to 1.5 (dependent on the interest rate series used) in Pagan et al. (1996).[9] For the 2-factor model, the estimate of λ is 0.59, which is very close to the 0.5 specification of the Cox-Ingersoll-Ross (1985a) type of volatility. The market prices of risk display very high rates of mean reversion. The mean reversion parameters α_1 and α_2 imply a halflife (the expected time that it takes for a state variable to return one half way back to its steady state level following a deviation) of around 0.5 and 0.7 months for the first and second market price of risk respectively.

Table 4 reports the prediction errors obtained by the models. It can be seen that the 2-factor model delivers the lowest mean absolute errors, averaging at 14 basis points across maturities. The errors are higher at the two ends of the yield curve. These prediction errors are in line with, and somewhat smaller than those in the study of Jegadeesh and Pennacchi (1996) who also used the Kalman filter (for a linear 2-factor model with constant volatilities) and reported a mean error (not mean absolute) of 23.5 basis points for the 1-month rate and 47.5 basis points for the 12-month rate.

Based on the estimates and the fitness of the models, the 2-factor model is our preferred choice. Figure 2 shows the volatilities of the short rate over the estimation period. The 2-factor model implies an average of 0.8% short rate volatility. The results for the period 1988-1992 are consistent with previous finding by Amin and Morton (1994), who studied the implied volatility of the short rate over that period. On average, the first volatility factor (the level effect) explains 99.87% whereas the second factor (the slope effect) explains 0.13% of the total variation of the yield curve.

As the volatility of the short rate increases, the drift should be lower (ie. at a discount compared to the zero volatility case), so that the corresponding drift for the bond price

[9] Bhar et al. (2005) have employed a Bayesian updating algorithm to estimate the distribution for the parameter λ in one factor HJM model implied by LIBOR rates of various maturities. They find that the distribution lies in the interval [0.5,4], giving support to the rather high (compared to some other studies) values for this parameter estimated in all markets here.

TABLE 3. Estimated Parameters for the U.S market

This table reports the estimated parameter values for each model using U.S data. The robust asymptotic standard errors are given in square parenthesis. "xe-y" stands for $x * 10^{-y}$

Par.	1-factor	2-factor	3-factor	Par.	1-factor	2-factor	3-factor
γ_1	3.2295 [3.83e-5]	0.0462 [3.33e-8]	1.2420 [4.26e-4]	α_1	11.4681 [8.33e-5]	15.9900 [1.19e-5]	6.2500 [0.0013]
γ_2	-	0.1261 [6.13e-9]	6.6221 [0.0023]	α_2	-	11.3091 [5.40e-7]	12.0068 [0.0006]
γ_3	-	-	0.0018 [4.89e-8]	α_3	-	-	11.7078 [0.0034]
λ	1.9662 [5.26e-6]	0.5859 [3.42e-7]	4.6823 [4.76e-5]	$\bar{\phi}_1$	18.5931 [1.47e-4]	39.5238 [1.75e-6]	18.6997 [0.0406]
κ	-	-	0.4623 [0.0001]	$\bar{\phi}_2$	-	16.2375 [1.45e-5]	0.5028 [1.56e-5]
σ_e^2	2.13e-6 [1.52e-11]	2.71e-7 [5.13e-13]	1.2e-6 [5.46e-9]	$\bar{\phi}_3$	-	-	19.6867 [0.0329]
				β_1	0.0104 [1.14e-7]	0.8189 [5.59e-9]	3.5873 [0.0092]
				β_2	-	4.1248 [1.78e-6]	0.0095 [4.14e-5]
$2 \ln \mathcal{L}$	33025.7	213025.8	28412.5	β_3	-	-	4.9948 [0.0007]

is higher which compensates investors for bearing higher risk. Figure 3 shows how these discounts (calculated by multiplying the standard deviation associated with each Wiener process by its corresponding market price of risk) are changing over time. As the level effect has a much larger impact on the volatility, most of the discounts are for this type of risk. The risk coming from the yield curve changing its slope is much lower, and therefore calls for a smaller bond premium.

4.2.2. *The Australian market.*

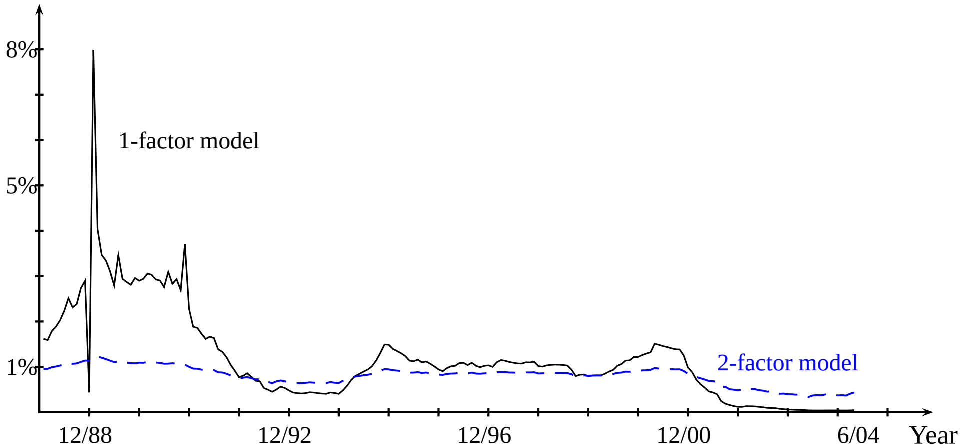
Similar to the U.S. market, the 2-factor model is also our preferred choice for the Australian market. The absolute prediction errors average at 30 basis points, which is slightly higher than that in the U.S. market. The estimate for λ is 3.5 times higher, predicting a smaller impact of the level of the interest rate on the overall volatility.

TABLE 4. The prediction errors for the U.S. market

This table reports the prediction errors for interbank offered rates of different maturities in the U.S. market. The "Avr" column reports the average of the prediction errors, whose standard deviation is reported under "Std" column. The "MAE" and "SAE" columns report the mean and standard deviation of the absolute errors. All values are reported as basis points.

	1-factor				2-factor				3-factor			
	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE
1-mth	-9.1	67.4	33.0	59.4	-19.5	25.3	26.4	18.0	36.7	38.2	45.0	27.9
2-mth	-5.8	63.2	27.8	57.0	-16.9	19.9	21.5	14.6	35.4	37.4	43.5	27.5
3-mth	-3.3	59.7	24.0	54.8	-15.1	15.4	18.2	11.6	33.6	37.3	42.1	27.2
4-mth	-1.4	57.0	20.4	53.2	-13.8	11.6	11.6	9.6	31.6	36.5	40.7	25.9
5-mth	0.9	54.8	18.1	51.7	-12.3	8.6	13.2	7.0	30.3	36.0	39.7	25.2
6-mth	2.7	52.9	17.6	49.9	-11.2	6.8	11.8	5.6	2.90	35.6	38.5	24.9
7-mth	5.3	52.1	19.3	48.6	-9.2	6.6	9.8	5.7	28.8	34.8	38.0	24.5
8-mth	7.5	51.5	21.3	47.5	-7.7	8.1	9.1	6.5	28.5	34.5	37.5	24.5
9-mth	9.9	51.6	23.8	46.8	-6.0	10.6	9.5	7.5	28.9	34.0	37.4	24.3
10-mth	12.6	52.3	26.8	46.7	-4.0	13.1	10.7	8.5	29.7	33.8	37.9	24.3
11-mth	15.6	53.4	30.3	46.6	-1.7	15.8	12.3	10.0	31.3	33.9	38.6	24.4
12-mth	18.4	55.1	33.8	47.1	0.4	18.6	14.5	11.6	32.9	33.4	39.7	24.9

FIGURE 2. The instantaneous volatilities of the U.S. short rate



The overall instantaneous short rate volatilities for the Australian market are graphed in Figure 4. Compared to the U.S. market, Australian rates have much higher volatility during the period 1988-1990, which reflects the very sharp rise and fall in the rates during that period. After 1990, the two markets have a similar volatility evolution. However, the contribution of each risk factor to this overall volatility is very different. In the U.S. market, the level factor explains more than 99.5% of the overall volatility

FIGURE 3. The “discount” on short rate drift to compensate for risk, 2-factor model, U.S. market

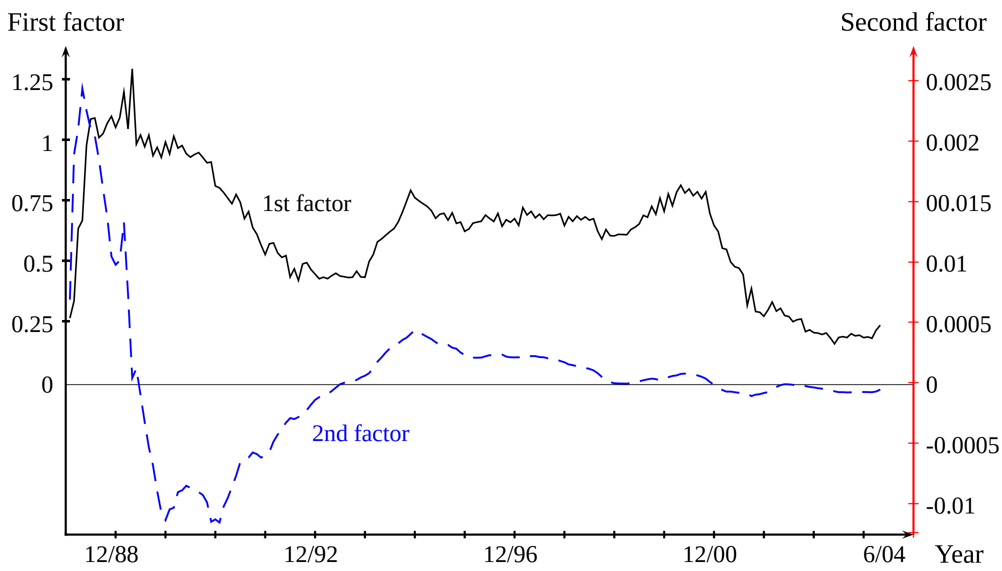
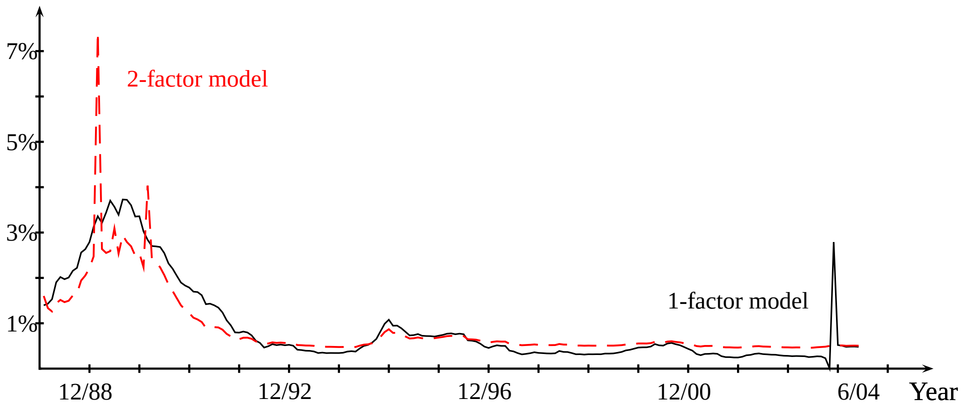


FIGURE 4. The instantaneous volatilities of the Australian short rate



throughout the whole period, whereas the slope effect plays a much more significant role in the Australian market, as can be seen in Figure 5.

Even though the slope factor contributes significantly to the overall interest rate volatility, the unit price of this risk is only a half of the unit price of the level risk. The long run value of ϕ_1 is 40 compared to the long run value of 19.9 for ϕ_2. The level of risk scaled by the unit price of risk is the discount on the short rate drift to compensate investors for bearing risk. Figure 6 shows how this discount changes over time. The

TABLE 5. Estimated Parameters for the Australian market

This table reports the estimated parameter values for each model using Australian data. The robust asymptotic standard errors are given in square parenthesis. "xe-y" stands for $x * 10^{-y}$

Par.	1-factor	2-factor	3-factor	Par.	1-factor	2-factor	3-factor
γ_1	0.9683 [3.22e-6]	0.9277 [4.52e-9]	1.2496 [0.0001]	α_1	9.3681 [1.42e-9]	9.1533 [2.26e-9]	8.9302 [0.3390]
γ_2	-	0.5322 [3.21e-9]	4.5897 [1.85e-7]	α_2	-	12.204 [3.54e-8]	12.4505 [0.0360]
γ_3	-	-	0.0187 [5.27e-5]	α_3	-	-	12.0037 [0.0298]
λ	1.8723 [1.20e-6]	1.9993 [9.69e-9]	4.9130 [0.0008]	$\bar{\phi}_1$	8.3133 [1.42e-9]	39.9988 [1.72e-7]	25.3901 [14.9103]
κ	-	-	0.0780 [0.0002]	$\bar{\phi}_2$	-	19.9088 [4.09e-7]	1.4157 [0.0004]
σ_e^2	2.14e-6 [4.66e-11]	3.34e-6 [1.87e-14]	2.59e-6 [5.90e-8]	$\bar{\phi}_3$	-	-	0.7040 [0.0030]
				β_1	0.1145 [4.53e-5]	0.0230 [4.50e-11]	5.9436 [1.7169]
				β_2	-	0.0337 [9.00e-11]	0.7964 [0.0021]
$2 \ln \mathcal{L}$	39198.9	49511.6	26626.4	β_3	-	-	0.7027 [0.0015]

FIGURE 5. The contribution of each factor toward the overall instantaneous volatility of the Australian short rate. 2-factor model.

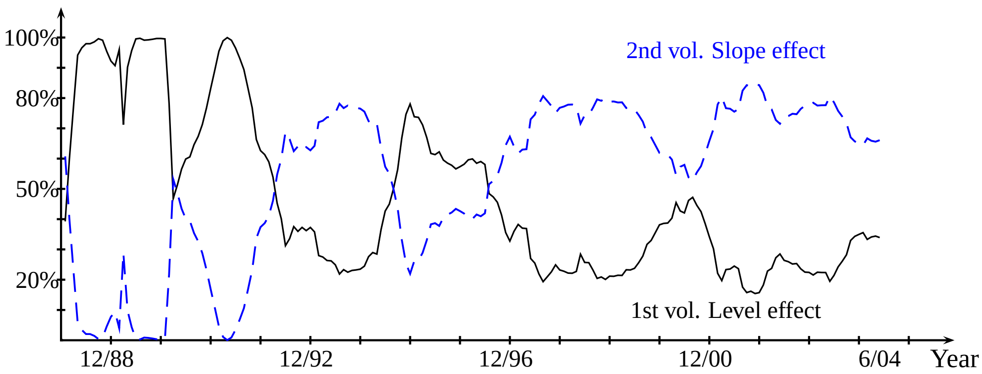
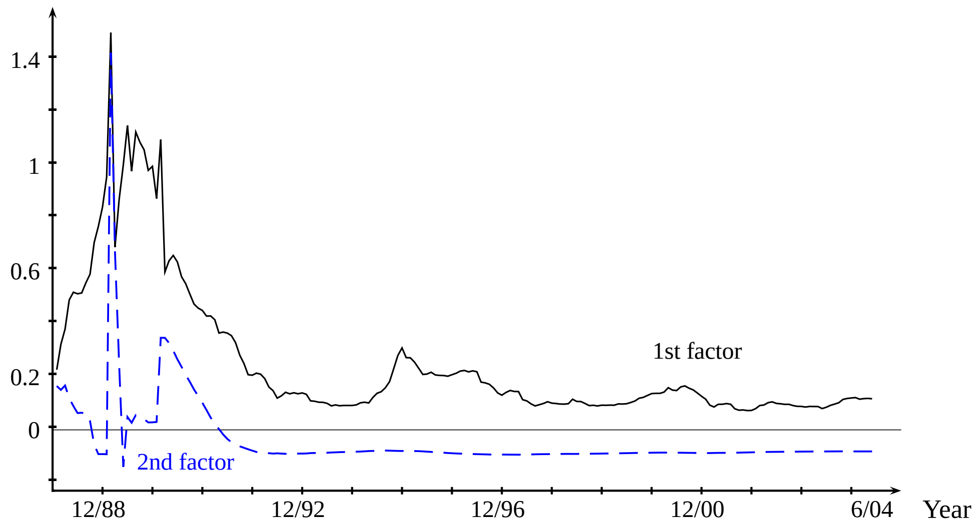


TABLE 6. The prediction errors for the Australian market

This table reports the prediction errors for interbank offered rate of different maturities in the Australian market. The “Avr” column reports the average of the prediction errors, whose standard deviation is reported under “Std” column. The “MAE” and “SAE” columns report the mean and standard deviation of the absolute errors. All values are reported as basis points.

	1-factor				2-factor				3-factor			
	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE
1-mth	-2.5	82.3	31.5	76.0	-13.2	96.8	40.2	89.0	3.0	119.9	83.2	86.3
2-mth	-0.9	79.8	26.9	75.1	-11.1	94.3	35.0	88.3	2.4	120.7	82.7	87.8
3-mth	-0.9	77.3	22.3	74.0	-10.6	93.2	29.2	89.2	0.4	120.9	82.0	88.6
4-mth	-0.5	75.3	18.3	73.1	-9.8	93.4	23.9	90.8	-1.2	120.9	81.4	81.4
5-mth	0.1	74.0	15.3	72.4	-8.7	94.3	19.3	92.7	-2.4	120.5	80.6	89.4
6-mth	1.2	73.1	14.2	71.7	-7.2	95.7	17.6	94.3	-3.0	119.5	79.5	89.1
7-mth	2.9	72.6	15.4	71.0	-5.0	98.5	19.7	96.6	-3.0	117.7	78.1	87.9
8-mth	4.6	72.5	18.2	70.3	-3.0	101.7	24.3	98.8	-2.9	115.2	76.5	86.1
9-mth	6.6	72.9	22.0	69.8	-0.5	105.4	29.9	101.1	-2.3	112.8	75.2	83.9
10-mth	8.8	73.6	25.6	69.5	2.1	109.3	35.7	103.3	-1.4	110.0	73.8	81.5
11-mth	11.0	74.6	29.4	69.4	4.8	113.8	42.0	105.8	-0.7	107.0	71.9	79.1
12-mth	13.1	75.7	32.9	69.3	7.3	118.1	47.8	108.2	0.6	103.2	70.0	75.5

FIGURE 6. The “discount” on short rate drift to compensate for risk, 2-factor model, Australian market



discount coming from bearing the risk of change in the level of interest rates is much higher than that coming from the risk of change in the slope of the yield curve.

4.2.3. *The U.K. market.*

Once again, the numerical optimization for the 3-factor model fails to distinguish the impact of each factor, see Table 7. The 2-factor model is also the preferred model in terms of model fitness. The prediction errors are below 40 basis points at the two end of the yield curve, and below 20 basis points in the middle range, as reported in Table 8. The estimates of λ and γ_2 are much higher than those in the U.S. or Australian market. Therefore, the contribution of each risk factor towards the total instantaneous volatility is the opposite of what happens in the U.S. market. Here the volatility coming from the slope risk factor dominates the total risk, and the level risk factor plays a negligible role. The total instantaneous short rate volatility, illustrated in Figure 7, has similar pattern as the Australian market, though at a slightly higher level.

TABLE 7. Estimated Parameters for the U.K. market

This table reports the estimated parameter values for each model using U.K. data. The robust asymptotic standard errors are given in square parenthesis. "xe-y" stands for $x * 10^{-y}$

Par.	1-factor	2-factor	3-factor	Par.	1-factor	2-factor	3-factor
γ_1	0.1171 [9.37e-5]	0.4560 [??]	1.1813 [0.0031]	α_1	11.5962 [0.0015]	11.6327 [??]	9.2382 [1.4875]
γ_2	-	9.9591 [??]	0.5843 [0.0014]	α_2	-	11.9505 [??]	12.4326 [0.0084]
γ_3	-	-	0.0163 [2.19e-5]	α_3	-	-	11.7322 [0.0010]
λ	1.0134 [0.0001]	3.8671 [??]	4.0522 [0.0008]	$\bar{\phi}_1$	6.8210 [0.0014]	37.8598 [??]	33.6271 [2.3820]
κ	-	-	0.0757 [5.99e-5]	$\bar{\phi}_2$	-	1.8489 [??]	5.7021 [0.1855]
σ_e^2	4.00e-6 [8.85e-9]	2.44e-6 [??]	3.80e-6 [2.85e-8]	$\bar{\phi}_3$	-	-	0.2516 [0.0047]
				β_1	0.1200 [1.38e-5]	4.1356 [??]	0.2293 [0.0034]
				β_2	-	4.6554 [??]	0.4558 [0.3468]
$2\ln\mathcal{L}$	34176.2	35599.6	26469.1	β_3	-	-	0.3697 [0.0009]

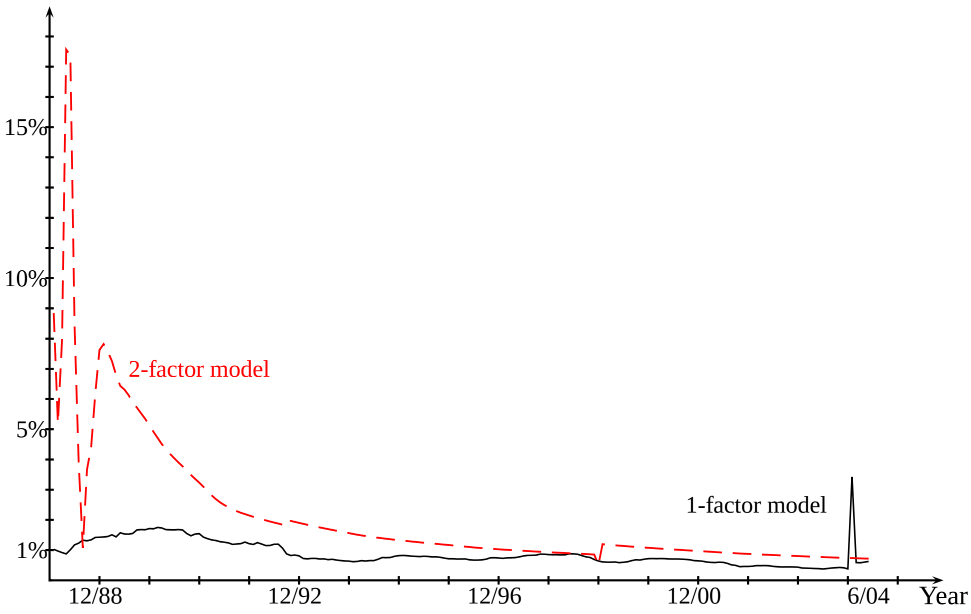
The unit price ϕ_1 of this level risk (the risk coming from changes in the level of interest rate) is of similar magnitude as in the other two markets. On the other hand,

TABLE 8. The prediction errors for the U.K. market

This table reports the prediction errors for interbank offered rate of different maturities in the U.K. market. The "Avr" column reports the average of the prediction errors, whose standard deviation is reported under "Std" column. The "MAE" and "SAE" columns report the mean and standard deviation of the absolute errors. All values are reported as basis points.

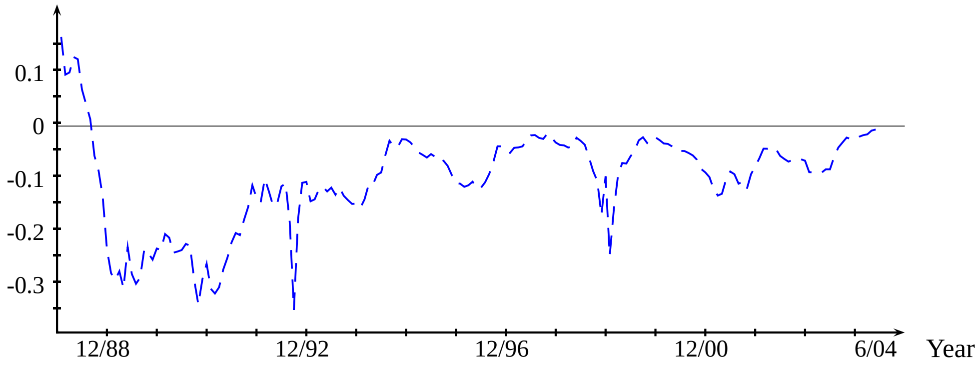
	1-factor				2-factor				3-factor			
	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE
1-mth	-21.2	184.8	40.8	181.5	-6.2	108.5	38.0	101.8	-6.7	70.1	48.7	50.7
2-mth	-16.8	182.4	33.5	180.1	-4.5	105.9	31.0	101.3	-4.5	70.5	47.1	52.6
3-mth	-14.1	180.6	28.4	178.9	-4.4	104.7	25.8	101.5	-3.9	71.0	46.0	54.1
4-mth	-12.5	178.9	23.5	177.8	-5.5	103.9	20.5	102.0	-4.3	70.1	44.5	54.2
5-mth	-10.8	177.6	19.4	176.9	-6.4	103.6	16.2	102.5	-4.4	69.4	43.6	54.2
6-mth	-9.2	176.5	18.4	175.8	-7.5	103.9	14.2	103.2	-4.7	68.8	42.9	54.0
7-mth	-7.1	175.6	20.4	174.6	-8.1	104.1	15.0	103.3	-4.4	67.2	41.9	52.8
8-mth	-5.1	175.0	24.3	173.3	-8.7	104.7	18.5	103.5	-4.0	66.4	41.6	51.9
9-mth	-3.1	174.5	28.6	172.2	-9.4	105.6	22.2	103.7	-3.5	65.5	41.5	50.7
10-mth	-0.6	174.2	32.8	171.1	-9.5	106.6	25.7	103.9	-2.5	64.5	41.2	49.5
11-mth	1.8	174.1	37.3	170.1	-9.8	107.9	29.5	104.2	-1.5	63.8	41.3	48.5
12-mth	4.2	174.1	41.5	169.1	-10.0	109.1	32.9	104.5	-0.4	63.1	41.6	47.3

FIGURE 7. The instantaneous volatilities of the U.K. short rate



the risk coming from the change in the slope of the yield curve calls for a much lower compensation. Despite this low compensation, due to the dominant risk value, almost

FIGURE 8. The "discount" on short rate drift to compensate for the slope risk, 2-factor model, U.K. market



all discount on the short rate drift to compensate investors for bearing risk is from the slope risk factor. The time variation nature of this discount can be seen in Figure 8.

4.2.4. *The Japanese market.*

Similar to other markets, the 2-factor model is the model that delivers the smallest absolute prediction errors in the Japanese market. The estimate for λ is 1.9, close to the level in the Australian market. Both of the market prices of risk have a high degree of mean reversion. One unit of the level risk is priced much more heavily than a unit of the slope risk, evidenced by the 35.9 estimate of $\bar{\phi}_1$ compare to the 4.5 estimate of $\bar{\phi}_2$.

At the end of 1995, the Japanese market moved to a period of low interest rates, slowly declining from around 50 basis points to around 5 basis points in 2002-2004. The instantaneous volatility of the short rate decreased accordingly, from a level of 30 basis points to nearly zero, as can be seen in Figure 9. During the low interest rate period, the factor that contributed most to interest rate risk was the slope of the yield curve. Figure 10 shows that the slope factor increased its influence throughout the declining period of 1991-1995, then became the most crucial risk factor during the near-zero interest rate of 1995-2004. However, each unit of slope risk claims less reward than one unit of the level risk, therefore the level risk still contributes significantly to the overall financial reward to investors, as illustrated by Figure 11.

TABLE 9. Estimated Parameters for the Japanese market

This table reports the estimated parameter values for each model using Japanese data. The robust asymptotic standard errors are given in square parenthesis. "xe-y" stands for $x * 10^{-y}$

Par.	1-factor	2-factor	3-factor	Par.	1-factor	2-factor	3-factor
γ_1	1.8982 [1.14e-7]	1.7255 [??]	1.4308 [0.0002]	α_1	13.9753 [8.34e-7]	12.4794 [??]	1.0739 [0.0032]
γ_2	-	1.1972 [??]	0.1249 [2.43e-6]	α_2	-	11.8275 [??]	11.3946 [0.0033]
γ_3	-	-	0.0097 [1.54e-6]	α_3	-	-	10.5193 [0.0005]
λ	1.9782 [5.21e-9]	1.8945 [??]	3.7487 [0.0009]	$\bar{\phi}_1$	35.3001 [4.56e-9]	35.8588 [??]	27.2719 [0.0192]
κ	-	-	0.0852 [2.76e-6]	$\bar{\phi}_2$	-	4.5297 [??]	12.4674 [0.0024]
σ_e^2	1.14e-7 [5.05e-15]	5.32e-7 [??]	1.13e-6 [5.94e-10]	$\bar{\phi}_3$	-	-	0.5566 [3.25e-5]
				β_1	0.0071 [3.40e-13]	0.0256 [??]	1.5606 [0.0013]
				β_2	-	0.0286 [??]	4.9166 [0.0063]
$2\ln\mathcal{L}$	43043.1	79896.8	29670.1	β_3	-	-	0.2134 [9.80e-5]

TABLE 10. The prediction errors for the Japanese market

This table reports the prediction errors for interbank offered rate of different maturities in the Japanese market. The "Avr" column reports the average of the prediction errors, whose standard deviation is reported under "Std" column. The "MAE" and "SAE" columns report the mean and standard deviation of the absolute errors. All values are reported as basis points.

	1-factor				2-factor				3-factor			
	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE	Avr	Std	MAE	SAE
1-mth	5.2	115.8	33.8	110.9	-7.3	63.1	19.3	60.5	3.6	30.3	17.5	25.0
2-mth	-19.3	138.4	24.6	137.6	-7.0	62.3	16.4	60.5	3.7	29.0	16.8	23.9
3-mth	-17.5	135.6	23.2	134.7	-6.9	62.5	14.1	61.3	3.6	29.0	16.2	24.4
4-mth	-15.8	132.1	21.7	131.2	-6.9	62.4	12.5	61.5	3.5	29.1	15.7	24.8
5-mth	-14.0	129.1	20.4	128.3	-6.8	62.7	11.3	62.1	3.4	29.0	15.5	24.7
6-mth	-11.6	126.6	18.6	125.7	-6.1	62.9	10.6	62.3	3.8	28.1	14.7	24.2
7-mth	-9.7	123.5	17.3	122.7	-5.9	63.5	10.5	62.9	3.9	28.0	14.6	24.2
8-mth	-7.4	121.0	17.1	120.0	-5.3	63.9	11.2	63.1	4.2	27.8	14.7	23.9
9-mth	-5.2	118.9	18.0	117.6	-4.1	64.8	12.6	63.8	4.6	27.8	15.3	23.6
10-mth	-2.8	116.4	19.2	114.8	-4.1	65.3	13.5	64.0	5.0	27.4	15.6	23.1
11-mth	-0.6	114.5	21.1	112.5	-3.6	66.2	14.8	64.6	5.3	27.5	16.3	22.7
12-mth	1.9	112.3	23.4	109.9	-2.8	66.7	15.9	64.8	5.9	27.4	27.4	22.3

FIGURE 9. The instantaneous volatilities of the Japanese short rate

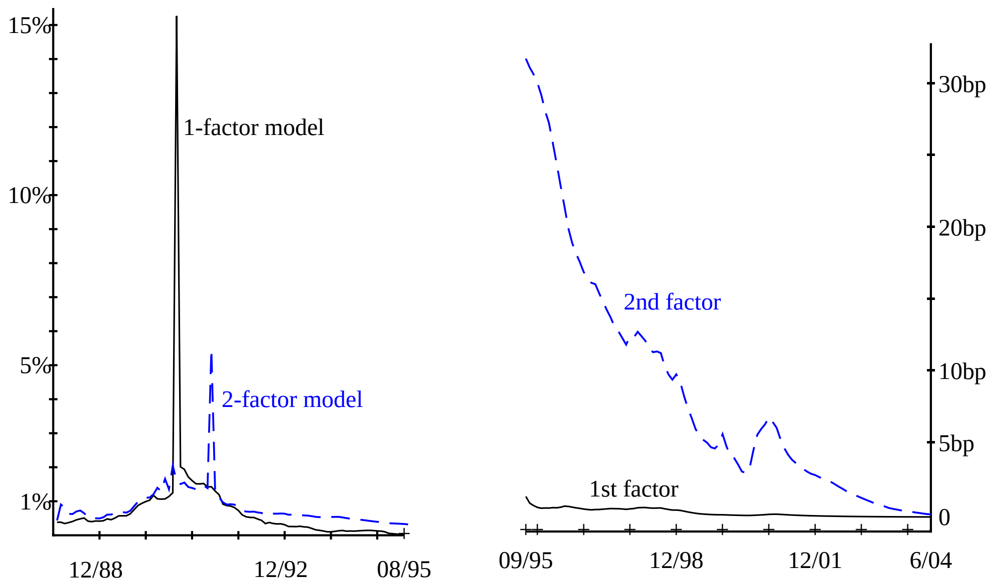


FIGURE 10. The contribution of each factor toward the overall instantaneous volatility of the Japanese short rate. 2-factor model.

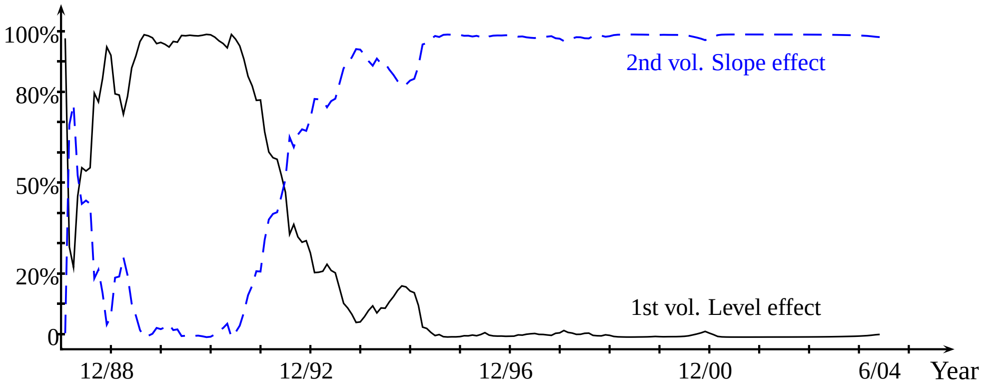
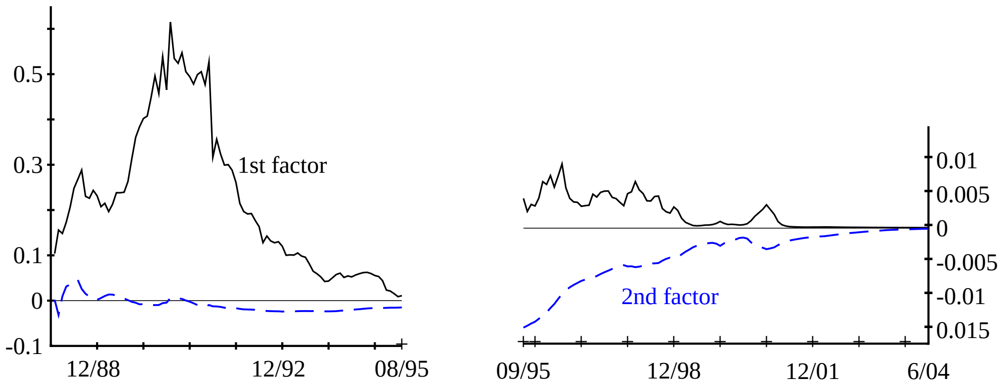


FIGURE 11. The "discount" on short rate drift to compensate for risk, 2-factor model, Japanese market



5. CONCLUSION

The HJM framework provides a very flexible tool for interest rate modelling. Even though theoretical research has advanced quickly, the advantages of HJM models have not been fully realized in practical applications due to the lack of empirical work. More research needs to be done on the challenging task of HJM model estimation in order to obtain a better understanding of interest rate volatility that is much needed in the process of assessing and managing risk as well as pricing derivative securities. This paper has attempted to contribute to the empirical literature by proposing an estimation framework that can be applied for a broad class of nonlinear HJM models.

The paper uses the local linearization filter to build up a maximum likelihood estimator which is able to identify all parameters of the model, and to exploit both time series and cross-sectional data. The local linearization scheme is based on an Itô-Taylor expansion of the nonlinear drift and diffusion terms of the driving dynamics to better take into account the stochastic behaviour of the interest rate system, and an optimal linear filter is subsequently applied. This filter has been chosen because of its advantages over other filters claimed by Shoji (1998) and its better numerical and stability properties demonstrated by Jimenez et al. (1999).

The estimator is then used to estimate the interest rate volatility structure in the U.S, the U.K, the Australian and the Japanese markets, using interbank offered-rates. In all markets, a 2-factor model, with the factors being the level and the slope of the yield curve, is found to be a reasonable choice. The influence of each factor on the overall instantaneous short rate volatility varies over time and across markets. The level factor is the dominant factor in the U.S market whereas the slope factor is the dominant one in the U.K. market. The two factors play a more equal role in the Australian market. In the Japanese market, the level effect has more impact on the overall volatility when interest rates are around a few percent, but the slope effect has more impact when interest rates stay at very low levels of less than 50 basis points.

Despite the different influence on the overall volatility, in all of the markets, the level risk claims a much higher financial reward than the slope risk. A knowledge of how each factor contributes to the overall volatility and the rewards for bearing the risk will help investors manage the risk of interest rate portfolios.

The filter adopted here is certainly not the only nonlinear filter available to modellers. It is left for future research to explore other filters, so as to find a good trade-off between reduction in computational requirements, increase in accuracy and better statistical reliability, all of which are crucial if financial managers are to re-assess their models frequently.

APPENDIX A. MARKOVIANIZATION OF THE INTEREST RATE DYNAMICS

Assuming that the forward rate $r(t, x)$ defined in (2.1) has a volatility function $\boldsymbol{\sigma}(t, x)$ that satisfies Assumption 2.1. Proposition 3.4 in Chiarella and Kwon (2003) states that the forward rate curve can be expressed as an affine function of some state variables, i.e.

$$\begin{aligned} r(t,x) = r(0, t+x) &+ \sum_{i=1}^{I} \sum_{l=1}^{L_i} \sigma_{il}(t+x)\psi_l^i(t) \\ &+ \sum_{i=1}^{I} \sum_{\substack{l,l^*=1 \\ l \le l^*}}^{L_i} [\sigma_{il}(t+x)\bar{\sigma}_{il^*}(t+x) + \epsilon_{ll^*}\sigma_{il^*}(t+x)\bar{\sigma}_{il}(t+x)]\varphi_{ll^*}^i(t), \end{aligned} \tag{A.1}$$

where

$$\begin{aligned} \bar{\sigma}_{il}(x) &= \int_0^x \sigma_{il}(s)\, ds, \\ \varphi_{ll^*}^i(t) &= \int_0^t c_{il}(s) c_{il^*}(s)\, ds, \\ \psi_l^i(t) &= \int_0^t c_{il}(s)\, d\widetilde{W}_i(s) - \sum_{l^*=1}^{d_i} \int_0^t c_{il}(s) c_{il^*}(s) \bar{\sigma}_{il^*}(s)\, ds, \\ \epsilon_{ll^*} &= \begin{cases} 1, & \text{if } l \neq l^*, \\ 0, & \text{if } l = l^*. \end{cases} \end{aligned}$$

and $\widetilde{W}_i, (i = 1, \ldots, I)$ are standard Wiener processes under the equivalent measure $\tilde{\mathcal{P}}$.

Under this setting, the economic meaning of the state variables φ and ψ is not clear. The next step is to use the forward rates themselves as the state variables.

Let $\mathscr{S} = \{\psi_l^i(t), \varphi_{lk}^i(t)\}$. Define $N = |\mathscr{S}|$, choose an ordering for $\mathscr{S}$ and write $\chi_n(t)$ for the elements of $\mathscr{S}$ so that $\mathscr{S} = \{\chi_1(t), \ldots, \chi_N(t)\}$. Then (A.1) can be written

$$r(t,x) = a_0(t,x) + \sum_{n=1}^{N} a_n(t,x)\, \chi_n(t), \tag{A.2}$$

for suitable deterministic functions $a_0(t, x)$ and $a_n(t, x)$.

Corollary A.1. *Suppose that the conditions of Assumption 2.1 are satisfied. If there exist* $\tau_1, \tau_2, \ldots, \tau_N \in \mathbb{R}_+$ *such that the matrix*

$$\boldsymbol{A}(t, \tau_1, \ldots, \tau_N) = \begin{bmatrix} a_1(t,\tau_1) & a_2(t,\tau_1) & \cdots & a_N(t,\tau_1) \\ a_1(t,\tau_2) & a_2(t,\tau_2) & \cdots & a_N(t,\tau_2) \\ \cdots & \cdots & \cdots & \cdots \\ a_1(t,\tau_N) & a_2(t,\tau_N) & \cdots & a_N(t,\tau_N) \end{bmatrix} \tag{A.3}$$

is invertible for all $t \in \mathbb{R}_+$*, then the variables* $\chi_n(t)$ *can be expressed in the form*

$$\boldsymbol{\chi}(t) = \boldsymbol{A}(t, \tau_1, \ldots, \tau_N)^{-1} \left[\boldsymbol{a}_0(t, \tau_1, \ldots, \tau_N) - \boldsymbol{r}(t, \tau_1, \ldots, \tau_N) \right], \tag{A.4}$$

where

$$\begin{aligned} \boldsymbol{\chi}(t) &= [\chi_1(t), \ldots, \chi_N(t)]', \\ \boldsymbol{a}_0(t, \tau_1, \ldots, \tau_N) &= [a_0(t,\tau_1), \ldots, a_0(t,\tau_N)]', \\ \boldsymbol{r}(t, \tau_1, \ldots, \tau_N) &= [r(t,\tau_1), \ldots, r(t,\tau_N)]'. \end{aligned}$$

The whole forward curve then can be written in terms of these new economically meaningful state variables

$$\begin{aligned} r(t,x) = a_0(t,x) &- \boldsymbol{a}(t,x)' \boldsymbol{A}(t, \tau_1, \ldots, \tau_N)^{-1} \boldsymbol{a}_0(t, \tau_1, \ldots, \tau_N) \\ &+ \boldsymbol{a}(t,x)' \boldsymbol{A}(t, \tau_1, \ldots, \tau_N)^{-1} \boldsymbol{r}(t, \tau_1, \ldots, \tau_N), \end{aligned} \tag{A.5}$$

where

$$\boldsymbol{a}(t,x) = [a_1(t,x), \ldots, a_N(t,x)]'.$$

Therefore, the HJM models admits a N-dimensional affine realization in terms of the set of discrete tenor forward rates $\boldsymbol{r}(t, \tau_1, \ldots, \tau_N)$. This set of forward rates forms a Markov process, and under $\widetilde{\mathcal{P}}$ each forward rate $r(t,x)$ satisfies the stochastic differential equation

$$\begin{aligned} dr(t,x) = &\left[\frac{\partial a_0(t,x)}{\partial x} - \frac{\partial \boldsymbol{a}(t,x)'}{\partial x} \boldsymbol{A}(t, \tau_1, \ldots, \tau_N)^{-1} \boldsymbol{a}_0(t, \tau_1, \ldots, \tau_N) \right. \\ &\left. + \frac{\partial \boldsymbol{a}(t,x)'}{\partial x} \boldsymbol{A}(t, \tau_1, \ldots, \tau_N)^{-1} \boldsymbol{r}(t, \tau_1, \ldots, \tau_N) + \boldsymbol{\sigma}(t, t+x)' \bar{\boldsymbol{\sigma}}(t, t+x) \right] dt \\ &+ \boldsymbol{\sigma}(t, t+x)' d\widetilde{\boldsymbol{W}}(t). \end{aligned}$$

In terms of the real world measure, where $\boldsymbol{\phi} \equiv (\phi_1, \ldots, \phi_I)$ is the vector of market prices of risk associating with the Wiener process $\boldsymbol{W}$, the system becomes

$$\begin{aligned} dr(t,x) = \Big[&\frac{\partial a_0(t,x)}{\partial x} - \frac{\partial \boldsymbol{a}(t,x)'}{\partial x} \boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1} \boldsymbol{a}_0(t,\tau_1,\ldots,\tau_N) \\ &+ \frac{\partial \boldsymbol{a}(t,x)'}{\partial x} \boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1} \boldsymbol{r}(t,\tau_1,\ldots,\tau_N) + \boldsymbol{\sigma}(t,t+x)' \bar{\boldsymbol{\sigma}}(t,t+x) \\ &- \boldsymbol{\phi}' \boldsymbol{\sigma}(t,t+x) \Big] dt + \boldsymbol{\sigma}(t,t+x)' d\boldsymbol{W}(t), \end{aligned}$$

which is (2.4) in the main text.

The yield $y(t,x)$ can also be expressed as an affine function of forward rates

$$\begin{aligned} y(t,x) = &b_0(t,x) - \boldsymbol{b}(t,x)' \boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1} \boldsymbol{a}_0(t,\tau_1,\ldots,\tau_N) \\ &+ \boldsymbol{b}(t,x)' \boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1} \boldsymbol{r}(t,\tau_1,\ldots,\tau_N), \end{aligned}$$

where

$$b_0(t,x) = \frac{1}{x} \int_0^x a_0(t,u) du,$$
$$\boldsymbol{b}(t,x) = \frac{1}{x} \int_0^x \boldsymbol{a}(t,u) du.$$

This affine yield expression is equation (2.6) in the main text.

Appendix B. Local Linearization Filter for Linear Continuous-Discrete State Space Models

Jimenez and Ozaki (2002) analyzed a linear state space model defined by the continuous state equation

$$d\mathbf{x}(t) = (\mathbf{A}(t)\mathbf{x}(t) + a(t))\, dt + \sum_{i=1}^{m} (\mathbf{B}_i(t)\mathbf{x}(t) + \mathbf{b}_i(t))\, d\mathbf{W}_i(t), \tag{B.1}$$

and the discrete observation equation[10]

$$\mathbf{z}_{t_j} = \mathbf{C}(t_j)\mathbf{x}(t_j) + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J, \tag{B.2}$$

where $\mathbf{x}(t) \in \mathbb{R}^d$ is the state vector at the instant of time t, $\mathbf{z}_{t_j} \in \mathbb{R}^r$ is the observation vector at the instant of time t_j, $\mathbf{W}$ is a m-dimensional vector of independent Wiener processes, and $\{\mathbf{e}_{t_j} : \mathbf{e}_{t_j} \sim \mathcal{N}(0, \mathbf{\Pi}), j = 0, \ldots, J\}$ is a sequence of random vector i.i.d.

[10] Their original specification is

$$\mathbf{z}_{t_j} = \mathbf{C}(t_j)\mathbf{x}(t_j) + \sum_{i=1}^{n} \mathbf{D}_i(t_j)\mathbf{x}(t_j)\boldsymbol{\xi}^i_{t_j} + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J,$$

where $\{\boldsymbol{\xi}_{t_j} : \boldsymbol{\xi}_{t_j} \sim \mathcal{N}(0, \mathbf{\Lambda}), \mathbf{\Lambda} = \mathrm{diag}((\lambda_1, \ldots, \lambda_n)), j = 0, \ldots, J\}$ is a sequence of random vector i.i.d., and $\mathbb{E}(\boldsymbol{\xi}^i_{t_j}, \mathbf{e}_{t_j}) = \boldsymbol{\vartheta}^i(t_j)$. However, in most finance applications, the noise term $\boldsymbol{\xi}$ is not required.

Define $\hat{\mathbf{x}}_{t|\rho} = \mathbb{E}\left(\mathbf{x}(t)|Z_\rho\right)$ and $\mathbf{P}_{t|\rho} = \mathbb{E}((\mathbf{x}(t) - \hat{\mathbf{x}}_{t|\rho})(\mathbf{x}(t) - \hat{\mathbf{x}}_{t|\rho})'|Z_\rho)$ for all $\rho \leq t$, where $Z_\rho = \{\mathbf{z}_{t_j} : t_j \leq \rho\}$.

Suppose that $\mathbb{E}\left(\mathbf{W}(t)\mathbf{W}'(t)\right) = \mathbf{I}$, $\hat{\mathbf{x}}_{t_0|t_0} < \infty$ and $\mathbf{P}_{t_0|t_0} < \infty$.

Theorem B.1. *(Jimenez and Ozaki (2002)) The optimal (minimum variance) linear filter for the linear model (B.1)- (B.2) consists of equations of evolution for the conditional mean $\hat{\mathbf{x}}_{t|t}$ and the covariance matrix $\mathbf{P}_{t|t}$. Between observations, these satisfy the ordinary differential equation*

$$d\hat{\mathbf{x}}_{t|t} = \left(\mathbf{A}(t)\hat{\mathbf{x}}_{t|t} + \mathbf{a}(t)\right) dt, \tag{B.3}$$

$$\begin{aligned} d\mathbf{P}_{t|t} = \Bigg(& \mathbf{A}(t)\mathbf{P}_{t|t} + \mathbf{P}_{t|t}\mathbf{A}'(t) + \sum_{i=1}^{m} \mathbf{B}_i(t) \left(\mathbf{P}_{t|t} + \hat{\mathbf{x}}_{t|t}\hat{\mathbf{x}}'_{t|t}\right) \mathbf{B}'_i(t) \\ & + \sum_{i=1}^{m} \left(\mathbf{B}_i(t)\hat{\mathbf{x}}_{t|t}\mathbf{b}'_i(t) + \mathbf{b}_i(t)\hat{\mathbf{x}}'_{t|t}\mathbf{B}'_i(t) + \mathbf{b}_i(t)\mathbf{b}'_i(t)\right) \Bigg) dt, \end{aligned} \tag{B.4}$$

for all $t \in [t_j, t_{j+1})$. At an observation at t_j, they satisfy the difference equation

$$\hat{\mathbf{x}}_{t_{j+1}|t_{j+1}} = \hat{\mathbf{x}}_{t_{j+1}|t_j} + \mathbf{K}_{t_{j+1}} \left(\mathbf{z}_{t_{j+1}} - \mathbf{C}(t_{j+1})\hat{\mathbf{x}}_{t_{j+1}|t_j}\right), \tag{B.5}$$

$$\mathbf{P}_{t_{j+1}|t_{j+1}} = \mathbf{P}_{t_{j+1}|t_j} - \mathbf{K}_{t_{j+1}}\mathbf{C}(t_{j+1})\mathbf{P}_{t_{j+1}|t_j}, \tag{B.6}$$

where

$$\mathbf{K}_{t_{j+1}} = \mathbf{P}_{t_{j+1}|t_j}\mathbf{C}'(t_{j+1})\left(\mathbf{C}(t_{j+1})\mathbf{P}_{t_{j+1}|t_j}\mathbf{C}'(t_{j+1}) + \mathbf{\Pi}\right)^{-1} \tag{B.7}$$

is the filter gain. The prediction $\hat{\mathbf{x}}_{t|\rho}$ and $\mathbf{P}_{t|\rho}$ are accomplished, respectively, via expressions (B.3) and (B.4) with initial conditions $\hat{\mathbf{x}}_{t_0|t_0}$ and $\mathbf{P}_{t_0|t_0}$ and $\rho < t$.

The analytical solution for this system of equations can be easily found, for details see Jimenez and Ozaki (2003). They also provide some equivalent expressions that are easier to implement via computer programs.

REFERENCES

Amin, K. I. and Morton, A. J. (1994), 'Implied Volatility Functions in Arbitrage-free Term Structure Models', *Journal of Financial Economics* **35**, 141–180.

Bhar, R. and Chiarella, C. (1997a), 'Transformation of Heath-Jarrow-Morton models to Markovian Systems', *European Journal of Finance* **3**, 1–26.

Bhar, R., Chiarella, C., Hung, H. and Runggaldier, W. (2005), The Volatility of the Instantaneous Spot Interest Rate Implied by Arbitrage Pricing - A Dynamic Bayesian Approach, Working paper, Quantitative Finance Research Centre, University of Technology, Sydney.

Björk, T. and Svensson, L. (2001), 'On the Existence of Finite Dimensional Realizations for Nonlinear Forward Rate Models', *Mathematical Finance* **11**, 205–243.

Bliss, R. R. and Ritchken, P. (1996), 'Empirical Tests of Two State-Variable Heath-Jarrow-Morton Models', *Journal of Money, Credit and Banking* **28**(3), 452–476.

Brace, A., Gatarek, D. and Musiela, M. (1997), 'The Market Model of Interest Rate Dynamics', *Mathematical Finance* **7**(2), 127–155.

Carverhill, A. (1994), 'When is the Short Rate Markovian?', *Mathematical Finance* **4**(4), 305–312.

Chan, K. C., Karolyi, G. A., Longstaff, F. A. and Sanders, A. B. (1992), 'An Empirical Comparison of Alternative Models of the Short-Term Interest Rate', *Journal of Finance* **47**, 1209–1221.

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

Chiarella, C. and Kwon, O. K. (2001b), 'Forward Rate Dependent Markovian Transformations of the Heath-Jarrow-Morton Term Structure Model', *Finance and Stochastics* **5**(2), 237–257.

Chiarella, C. and Kwon, O. K. (2003), 'Finite Dimensional Affine Realisations of HJM Models in Terms of Forward Rates and Yields', *Review of Derivatives Research* **6**(2), 129–155.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985a), 'An Intertemporal General Equilibrium Model of Asset Prices', *Econometrica* **53**(2), 363–384.

de Jong, F. and 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.

Duffee, G. R. and Stanton, R. H. (2004), Estimation of Dynamic Term Structure Models, Working paper, Haas School of Business, U.C. Berkeley.

Duffie, D. and Kan, R. (1996), 'A Yield-Factor Model of Interest Rates', *Mathematical Finance* **6**(4), 379–406.

Heath, D., Jarrow, R. and Morton, A. (1992), 'Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation', *Econometrica* **60**(1), 77–105.

Holland, J. (1975), *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor.

Inui, K. and Kijima, M. (1998), 'A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models', *Journal of Financial and Quantitative Analysis* **33**(3), 423–440.

Jegadeesh, N. and Pennacchi, G. (1996), 'The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures', *Journal of Money, Credit and Banking* **28**(3), 426–446.

Jimenez, J. C. and Ozaki, T. (2002), 'Linear Estimation of Continuous-Discrete Linear State Space Models with Multiplicative Noise', *Systems and Control Letters* **47**, 91–101. Erratum appears in V49, page 161, 2003.

Jimenez, J. C. and Ozaki, T. (2003), 'Local Linearization Filters for Non-Linear Continuous-Discrete State Space Models with Multiplicative Noise', *International Journal of Control* **76**(12), 1159–1170.

Jimenez, J. C., Shoji, I. and Ozaki, T. (1999), 'Simulation of Stochastic Differential Equations through the Local Linearization Method. A Comparative Study', *Journal of Statistical Physics* **94**, 587–602.

Knez, P. J., Litterman, R. and Scheinkman, J. (1994), 'Exploration into Factors Explaining Money Market Returns', *Journal of Finance* **49**(5), 1861–1882.

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

Michalewicz, Z. (1999), *Genetic Algorithms + Data Structures = Evolution Programs*, Springer-Verlag.

Mitchell, M. (1996), *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA.

Pagan, A. R., Hall, A. D. and Martin, V. (1996), Modeling the Term Structure, *in* G. S. Maddala and C. R. Rao, eds, 'Handbook of Statistics', Vol. 14, Elsevier Science, chapter 4, pp. 91–118.

Ritchken, P. and Sankarasubramanian, L. (1995), 'The Importance of Forward Rate Volatility Structures in Pricng Interest Rate-Sensitive Claims', *Journal of Derivatives* **3**(1), 25–41.

Shoji, I. (1998), 'A Comparative Study of Maximum Likelihood Estimators for Nonlinear Dynamical System Models', *International Journal of Control* **71**(3), 391–404.

Singh, M. K. (1995), 'Estimation of Multifactor Cox, Ingersoll, and Ross Term Structure Model: Evidence on Volatility Structure and Parameter Stability', *Journal of Fixed Income* **5**(2), 8–28.

Vose, M. D. (1999), *The Simple Genetic Algorithm: Foundations and Theory*, MIT Press, Cambridge, MA.