THE VOLATILITY STRUCTURE OF THE FIXED INCOME MARKET UNDER THE HJM FRAMEWORK

CARL CHIARELLA* AND THUY-DUONG TÔ**

School of Finance and Economics
University of Technology, Sydney
PO Box 123
Broadway NSW 2007
Australia
*carl.chiarella@uts.edu.au
**td.to@uts.edu.au

ABSTRACT. This paper considers the dynamics for interest rate processes within the Heath, Jarrow and Morton (1992) specification. It is well known that one of the difficulties in using this specification for estimation is the non-Markovian nature of the dynamics. The paper focuses on a fairly broad family of models that not only can be transformed into a Markovian dynamics, but also has an affine representation for the observed data, which overlaps but is not nested in the Duffie and Kan (1996) class of affine term structure models. The model parameters are estimated using a maximum likelihood function obtained via the local linearization filter proposed by Jimenez and Ozaki (2002, 2003). The method is then applied to analyze the volatility structure of the LIBOR markets.

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

JEL classifications: C51, E43, G12

1. INTRODUCTION

The no-arbitrage approach, as opposed to the general equilibrium approach, to interest rate modelling has wide appealing power. The specification of Heath, Jarrow and Morton (1992) (hereafter HJM), and later of Brace, Gatarek and Musiela Brace et al. (1997) have proved the flexibility of the approach in capturing the shape of the yield curve and the ability to price various financial quantities. Even though there has been huge advancement in the theoretical field, the practical implementation of this

---

Date: Revision: May 31, 2004.
Preliminary and incomplete. Please do not quote.
** Corresponding author.

model framework is hindered by the non-markovian nature of the interest rate dynamics. Estimation of the model parameters becomes a challenging problem despite its obvious importance in risk management, pricing and forecast activities.

This paper considers the class of the HJM models analyzed by Chiarella and Kwon (2001b, 2003), where the interest rate dynamics can be markovianized (by extending the state variable vector) and the observable financial quantities such as the yields or bond prices can be expressed as affine functions of the underlying state variables. This class nested some important classes that have been considered previously in the HJM framework, such as those analyzed by Carverhill (1994), Ritchken and Sankarasubramanian (1995), Bhar and Chiarella (1997a), Inui and Kijima (1998) and De Jong and Santa-Clara (1999). In addition, this class of models is not nested in the affine term structure class considered in Duffie and Kan (1996), even though there will be occasions when the two classes overlap.

Even though a Markovian structure is obtained, the resulting interest rate system is high-dimensional and nonlinear, imposing difficulty in estimation. This paper advocates the local linearization method proposed by Jimenez and Ozaki (2002, 2003) to calculate the likelihood function of observable data. The idea is to linearize the stochastic differential equations of the state variables via Itô formula to take into account their stochastic behaviours. The linear stochastic differential equations are then solved analytically for their discrete representations and a simple Kalman filter can be called to calculate the likelihood function for the observables and therefore the model can be estimated via the maximum likelihood method.

The method is then applied to analyze the volatility structure of LIBOR rates traded on the United States (hereafter U.S.), the Great Britain, the Australian and the Japanese markets. The rest of the paper is structured as follows. Section 2 discusses the class of models we are looking at. Section 3 analyzes the econometric implications of the model, what has been done in the literature and our proposal. Empirical analysis is carried out in Section 4 and finally 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 confracted at t for instantaneous borrowing/lending at $t+x$)

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 Wiener process 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 $d\boldsymbol{W}(t)$.

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. The research of Carverhill (1994), Ritchken and Sankarasubramanian (1995), Bhar and Chiarella (1997a), Inui and Kijima (1998), De Jong and Santa-Clara (1999) has advanced our knowledge on the field. Recently, Chiarella and Kwon(2001b, 2003) introduce a class of models that can specialize into the above-mentioned 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[2] that the forward curve can be expressed as an affine function of some state variables,

[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] See proposition 3.4 in their paper.

Proposition 2.2. *Let $\sigma(t,x)$ satisfy the conditions of Assumption 2.1. Then the forward rate curve can be expressed in the form*

$$r(t,x) = r(0,t+x) + \sum_{i=1}^{I}\sum_{l=1}^{L_i} \sigma_{il}(t+x)\psi_l^i(t) + \sum_{i=1}^{I}\sum_{\substack{l,l^*=1 \\ l\le l^*}}^{L_i} [\sigma_{il}(t+x)\bar{\sigma}_{il^*}(t+x) + \epsilon_{ll^*}\sigma_{il^*}(t+x)\bar{\sigma}_{il}(t+x)]\varphi_{ll^*}^i(t), \tag{2.4}$$

where

$$\bar{\sigma}_{il}(x) = \int_0^x \sigma_{il}(s)\,ds, \tag{2.5}$$

$$\varphi_{ll^*}^i(t) = \int_0^t c_{il}(s)c_{il^*}(s)\,ds, \tag{2.6}$$

$$\psi_l^i(t) = \int_0^t c_{il}(s)\,d\widetilde{W}_i(s) - \sum_{l^*=1}^{d_i}\int_0^t c_{il}(s)c_{il^*}(s)\bar{\sigma}_{il^*}(s)\,ds, \tag{2.7}$$

$$\epsilon_{ll^*} = \begin{cases} 1, & \textit{if } l \neq l^*, \\ 0, & \textit{if } l = l^*. \end{cases} \tag{2.8}$$

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

Let $y(t,x)$ be the yield on the $(t+x)$-maturity zero coupon bond $b(t,x)$. These quantities are uniquely linked with the instantaneous forward rates via,

$$b(t,x) = \exp\left(-\int_0^x r(t,u)du\right), \tag{2.9}$$

$$y(t,x) = \frac{1}{x}\int_0^x r(t,u)du, \tag{2.10}$$

therefore they can also be expressed as an affine function of the state variables.

The functional forms for the state variables indicate that this class of model is not nested in the affine term structure class considered in Duffie and Kan (1996), even though there will be occasions when the two classes overlap. Since it is unclear what these state variables represent, the forward/yield curve can be "inverted" so that economically meaningful quantities can serve 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 $\mathscr{S} = \{\chi_1(t),\dots,\chi_N(t)\}$. Then (2.4) can be written

$$r(t,x) = a_0(t,x) + \sum_{n=1}^{N} a_n(t,x)\,\chi_n(t), \tag{2.11}$$

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

Corollary 2.3. *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*

$$\boldsymbol{A}(t,\tau_1,\ldots,\tau_N) = \begin{bmatrix} a_1(t,\tau_1) & a_2(t,\tau_1) & \cdots & a_N(t,\tau_1) \\ a_1(t,\tau_2) & a_2(t,\tau_2) & \cdots & a_N(t,\tau_2) \\ \cdots & \cdots & \cdots & \cdots \\ a_1(t,\tau_N) & a_2(t,\tau_N) & \cdots & a_N(t,\tau_N) \end{bmatrix} \tag{2.12}$$

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

$$\boldsymbol{\chi}(t) = \boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1}\left[\boldsymbol{a}_0(t,\tau_1,\ldots,\tau_N) - \boldsymbol{r}(t,\tau_1,\ldots,\tau_N)\right], \tag{2.13}$$

where

$$\boldsymbol{\chi}(t) = [\chi_1(t),\ldots,\chi_N(t)]', \tag{2.14}$$

$$\boldsymbol{a}_0(t,\tau_1,\ldots,\tau_N) = [a_0(t,\tau_1),\ldots,a_0(t,\tau_N)]', \tag{2.15}$$

$$\boldsymbol{r}(t,\tau_1,\ldots,\tau_N) = [r(t,\tau_1),\ldots,r(t,\tau_N)]'. \tag{2.16}$$

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

$$\begin{aligned} r(t,x) = {} & a_0(t,x) - \boldsymbol{a}(t,x)'\boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1}\boldsymbol{a}_0(t,\tau_1,\ldots,\tau_N) \\ & + \boldsymbol{a}(t,x)'\boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1}\boldsymbol{r}(t,\tau_1,\ldots,\tau_N), \end{aligned} \tag{2.17}$$

where

$$\boldsymbol{a}(t,x) = [a_1(t,x),\ldots,a_N(t,x)]'. \tag{2.18}$$

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

$$\begin{aligned} dr(t,x) = {} & \left[\frac{\partial a_0(t,x)}{\partial x} - \frac{\partial \boldsymbol{a}(t,x)'}{\partial x}\boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1}\boldsymbol{a}_0(t,\tau_1,\ldots,\tau_N)\right. \\ & \left. + \frac{\partial \boldsymbol{a}(t,x)'}{\partial x}\boldsymbol{A}(t,\tau_1,\ldots,\tau_N)^{-1}\boldsymbol{r}(t,\tau_1,\ldots,\tau_N) + \boldsymbol{\sigma}(t,t+x)'\bar{\boldsymbol{\sigma}}(t,t+x)\right] dt \\ & + \boldsymbol{\sigma}(t,t+x)'d\widetilde{\boldsymbol{W}}(t). \end{aligned} \tag{2.19}$$

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

$$\begin{aligned} dr(t,x) = \Big[& \frac{\partial a_0(t,x)}{\partial x} - \frac{\partial \boldsymbol{a}(t,x)'}{\partial x} \boldsymbol{A}(t,\tau_1,\dots,\tau_N)^{-1} \boldsymbol{a}_0(t,\tau_1,\dots,\tau_N) \\ & + \frac{\partial \boldsymbol{a}(t,x)'}{\partial x} \boldsymbol{A}(t,\tau_1,\dots,\tau_N)^{-1} \boldsymbol{r}(t,\tau_1,\dots,\tau_N) + \boldsymbol{\sigma}(t,t+x)' \bar{\boldsymbol{\sigma}}(t,t+x) \\ & + \boldsymbol{\phi}' \boldsymbol{\sigma}(t,t+x) \Big] dt + \boldsymbol{\sigma}(t,t+x)' d\widetilde{\boldsymbol{W}}(t). \end{aligned} \tag{2.20}$$

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

$$\begin{aligned} y(t,x) = & b_0(t,x) - \boldsymbol{b}(t,x)' \boldsymbol{A}(t,\tau_1,\dots,\tau_N)^{-1} \boldsymbol{a}_0(t,\tau_1,\dots,\tau_N) \\ & + \boldsymbol{b}(t,x)' \boldsymbol{A}(t,\tau_1,\dots,\tau_N)^{-1} \boldsymbol{r}(t,\tau_1,\dots,\tau_N), \end{aligned} \tag{2.21}$$

where

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

3. ESTIMATION FRAMEWORK

3.1. **Econometric implication of the model.**

Some specialized models of the class discussed in the previous section have been empirically analyzed. Bliss and Ritchken (1996) consider the case where the volatility function in (2.2) can be written as[3]

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

They then introduce an error term into the relationship (2.21) to estimate their model via the Maximum Likelihood procedure [4]. This method has two undesirable features. First, due to the need to input the market values of the forward rates into (2.21), the estimation results may vary as to the choice of which forward rates to serve as the state variables. This choice is arbitrary, and Bliss and Ritchken also show that the parameter estimates are sensitive to this choice. Second, Bliss and Ritchken realize that the relationship (2.21) does not depend on the parameters characterizing function $c(t)$, and therefore their estimation method can only identify κ [5]. However, all of the

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

[4]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.

[5]In our general class, the only parameters that affect the relationship (2.21) are those of the functions σ_{ij}. Those parameters of the function c_{ij} do not affect this linkage between term structure.

parameters in the models are important in practical work, such as determining the price of a derivative contract.

De Jong and Santa-Clara (1999) also empirically study two-state variable HJM model. However, they overcome the problem of Bliss and Ritchken (1996) by using both the relationship (2.21) and the markovian system (2.20) in their estimation. They use the Kalman filtering method where (2.21) serves as the observable equation and (2.20) is discretized into a transition equation. The market price of risk term in (2.20) is assumed to be proportional to the square root of the spot rate. In their paper, due to the special structure of their model, De Jong and Santa-Clara can solve the system for an exact linear discretized solution to apply the Kalman filter. However, not all nonlinear system has an exact solution and the behaviour of the estimator, is therefore, dependent on the method of discretizing the structural stochastic system.

In this paper, we advocate the local linearization filter (hereafter LL filter) of Jimenez and Ozaki (2002, 2003). The approach is still based on the Kalman filter for a discrete linear system. However, Jimenez and Ozaki do not discretize the nonliear system directly, but rather approximate it by a linear system, then solve the linear system for an exact solution, and finally apply the Kalman filter. The approximation is not based on a first order Taylor approximation used in the extended Kalman filter framework, but it is based on the second order approximation using Ito formula to 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 variable). Shoji, via Monte Carlo method, showed that the LL filter provided estimates with smaller bias, particularly in estimation of the coefficient of the drift term. In a different study, Jimenez et al. (1999) also reported the numerical advantages of the LL filter, including numerical stability, better accuracy and order of strong convergence.

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

Let the state space model defined by the continuous state equation

$$d\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t))dt + \sum_{i=1}^{m} \mathbf{g}_i(t, \mathbf{x}(t))dW_i(t), \tag{3.1}$$

and the discrete observation equation[6]

$$\mathbf{z}_{t_j} = \mathbf{C}(t_j)\mathbf{x}(t_j) + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J, \tag{3.2}$$

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 a 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 random vector i.i.d.

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

$$\begin{aligned}\mathbf{f}(t, \mathbf{x}(t)) \approx & f(s, \mathbf{u}) + \left(\frac{\partial \mathbf{f}(s, \mathbf{u})}{\partial s} + \frac{1}{2} \sum_{k,l=1}^{d} [\mathbf{G}(s, \mathbf{u})\mathbf{G}'(s, \mathbf{u})]^{k,l} \frac{\partial^2 \mathbf{f}(s, \mathbf{u})}{\partial \mathbf{u}^k \partial \mathbf{u}^l} \right)(t - s) \\ & + \mathbf{J_f}(s, \mathbf{u})(\mathbf{x}(t) - \mathbf{u}),\end{aligned} \tag{3.3}$$

where $(s, \mathbf{u}) \in \mathbb{R} \times \mathbb{R}^d$, $\mathbf{J_f}(s, \mathbf{u})$ 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 these approximations for $\mathbf{f}$ and $\mathbf{g}_i$, the solution of the nonlinear state equation (3.1) can be approximated by the solution of the piecewise linear sde

$$\begin{aligned}d\mathbf{y}(t) = & \left(\mathbf{A}(t_j, \hat{\mathbf{y}}_{t_j|t_j})\mathbf{y}(t) + \mathbf{a}(t, t_j, \hat{\mathbf{y}}_{t_j|t_j}) \right) dt \\ & + \sum_{i=1}^{m} \left(\mathbf{B}_i(t_j, \hat{\mathbf{y}}_{t_j|t_j})\mathbf{y}(t) + \mathbf{b}_i(t, t_j, \hat{\mathbf{y}}_{t_j|t_j}) \right) dW_i(t)\end{aligned} \tag{3.4}$$

[6]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 nonliear 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 view of most finance applications, including ours, we chose a linear specification for $\mathbf{p}_i$ and to omit the term $\boldsymbol{\xi}$.

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 remaining notations are defined as

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

This linear state equation can be solved for an exact solution.

After approximation, (3.4) and the "new" observation equation (see (3.2))

$$\mathbf{z}_{t_j} = \mathbf{C}(t_j)\mathbf{y}(t_j) + \mathbf{e}_{t_j}, \text{ for } j = 0, 1, \ldots, J, \tag{3.5}$$

form a linear state space system. The optimal linear filter proposed by Jimenez and Ozaki (2002) can be applied (see Appendix A for the 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 $\rho \leq t$.

Due to the assumption of multivariate normality of the disturbances $\mathbf{e}_{t_j}$ (and if the initial state vector also have a proper multivariate normal distribution), the distribution of $\mathbf{z}_{t_{j+1}}$ conditional on Z_{t_j} is itself normal (see (3.5)). 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.

Let $\boldsymbol{\theta}$ be the vector of parameters of interest, which include all parameters specifying the state space model (3.4) and (3.5), 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

$$\mathcal{L}_{\mathbf{Z}}(\boldsymbol{\theta}) = -\frac{rJ}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{J}\ln|\boldsymbol{\Sigma}_{t_j}| - \frac{1}{2}\sum_{j=1}^{J}\boldsymbol{\nu}'_{t_j}\boldsymbol{\Sigma}_{t_j}^{-1}\boldsymbol{\nu}_{t_j} \tag{3.6}$$

where the innovation equations are

$$\boldsymbol{\nu}_{t_j} = \mathbf{z}_{t_j} - \mathbf{C}(t_j)\hat{\mathbf{y}}_{t_j|t_{j-1}}, \tag{3.7}$$

$$\boldsymbol{\Sigma}_{t_j} = \mathbf{C}(t_j)\mathbf{P}_{t_j|t_{j-1}}\mathbf{C}'(t_j) + \boldsymbol{\Pi}. \tag{3.8}$$

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

$$\hat{\boldsymbol{\theta}} = \max_{\boldsymbol{\theta}} \mathcal{L}_{\mathbf{Z}}(\boldsymbol{\theta}). \tag{3.9}$$

3.3. **Econometric implementation.**

We analyze a one-factor interest rate model where the volatility of the interest rates is dependent on the level of the rates, i.e.

$$\sigma(t,x) = \gamma r^{\lambda}(t). \tag{3.10}$$

The dependence of interest rate volatility on the level of rates has been long documented. In a large number of previous research, especially within the quadratic term structure family of models, a square root dependence (i.e. $\lambda = 0.5$) has been used. However, some other researchers have estimated λ and found it to be equal to 1.5 in Chan et al. (1992), and in the range of 0.5 to 1.5 (dependent on the interest rate series used) in Pagan et al. (1996).

The market price of risk terms ϕ is assumed to follow a CIR type of processes, ie. it is mean reverting and has volatility function proportional to the square root of the level of itself, ie.

$$d\phi = \alpha(\bar{\phi} - \phi)dt + \beta\sqrt{\phi(t)}dW(t). \tag{3.11}$$

Intuitively, the specification suggests that the market price of interest rate risk is always positive and tends to converge to its long run equilibrium.

Using some algebra work, we now have a continuous-discrete nonlinear state space system. The nonlinear transitional system contains the evolution equations for three underlying state variables

$$dr(t) = \left(\frac{r(t,\tau_2) - r(t)}{\tau_2} - \phi\gamma r^{\lambda}\right) + \gamma r^{\lambda} dW(t), \tag{3.12}$$

$$dr(t,\tau_2) = \left(\frac{r(t,\tau_2) - r(t)}{\tau_2} + \gamma^2 r^{2\lambda}\tau_2 - \phi\gamma r^{\lambda}\right) + \gamma r^{\lambda} dW(t), \tag{3.13}$$

$$d\phi(t) = \alpha(\bar{\phi} - \phi)dt + \beta\sqrt{\phi}dW(t). \tag{3.14}$$

The linear measurement equation expresses zero-coupon yields as functions of the state variables

$$y(t,x) = \left(1 - \frac{x}{2\tau_2}\right) r(t) + \frac{x}{2\tau_2} r(t,\tau_2) + e_x \tag{3.15}$$

(for different values of maturity x), where we have introduced 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 LL algorithm can be readily implemented. We use LIBOR rates for our empirical study. It should be noted that there is a one-to-one relationship between Libor rate and zero-coupon yield

$$\exp(x \times y(t, x)) = 1 + xL(t, x), \tag{3.16}$$

therefore to obtain the likelihood function for these Libor series we need to transform the likelihood function for the yields obtained through the LL procedure via a Jacobian transformation.

4. Empirical Analysis

4.1. **The Data.**

We use weekly data for Libor rate with maturities range from 1 month to 12 months traded during the 4-year period 1999-2002 in 4 markets of the United States, the Great Britain, the Japan and the Australia.

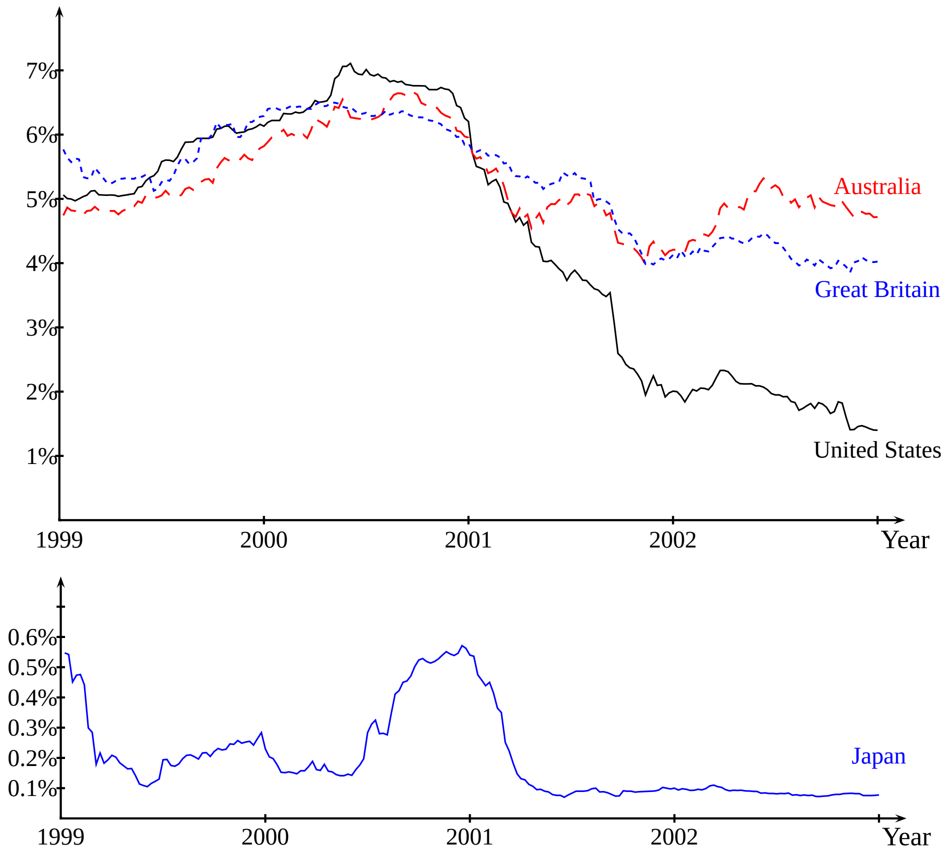
The typical movements of the LIBOR rates in each market can be seen in Figure 1. In Australia and Great Britain, the rates fluctuated around a mean level of 5.3%. At the end of year 1999, the rate in the Great Britain started to decline and kept this movement throughout the year 2000. The Australian market experienced a much shaper decline in the last quarter of 1999 and first quarter of 2000, and then moved widely around the mean level. The U.S. market experienced even a shaper decline until the beginning of 2002, when the rates gradually decreased from 2% to a low level of 1.5%. However, this level was still much higher than that in the Japanese market. This market had an extremely high volatility of very low interest rate level, which was as low as 0.1% since the second quarter of 2001.

4.2. **Empirical Results.**

The model (3.12)-(3.15) are fitted to the four markets, and the estimation results are reported in Table 1.

The dependence of the volatility on the level of interest rate is a nonlinear relationship via the power parameter λ. Only in the Japanese market where the estimated λ is 0.47 we find evidence of a square-root typed dependency. In all of other markets we find an estimate of 1.5, which predicts a much smaller volatility for a given level of interest rate. The implied volatilities for the instantaneous rate vary widely

FIGURE 1. 6-month LIBOR Rates



between different markets (see Figure 2). The Japanese market, with its extremely low interest rate level, has the highest implied volatility, averaging at an astonishing level of 15%. The Great Britain and Australia markets have stable interest rate volatility, though volatility in the Australian market is averaging at 5%, which is much lower than the level of 7.5% in the Great Britain. In the U.S., the volatility picks at the beginning of 2001, when the sharp decline in the level of rate is started. The volatility level fluctuates during this rate-declining period then gradually settles down to the 1% level.

Figure 3 illustrates the implied market price of interest rate risk in each market[7]. The U.S and the Australian market offer a highest reward for a given level of interest rate risk. The Japanese market pays the smallest price for the risk and the Great Britain only reward a slightly higher prize compared to the Japanese market. All of the market prices of risk revert very quickly towards their long term level. It takes roughly 1.7

[7]The first 3 observations are ignored. The values which seem to be unreasonable might be the result of the filter has not settled down properly.

TABLE 1. Estimation Results

This table reports the parameter estimates for each market. Asymptotic standard errors of the estimate are given in parentheses.

Parameter	United States	Australia	Great Britain	Japan
γ	3.7499	3.0127	4.6875	3.7499
	(0.0041)	(1.73e-9)	(6.16e-5)	(8.88e-9)
λ	1.4963	1.4989	1.4999	0.4683
	(0.0009)	(3.42e-7)	(1.10e-7)	(1.92e-8)
α	49.9983	49.1699	99.9023	98.4375
	(0.0019)	(4.85e-7)	(1.14e-5)	(1.22e-8)
$\bar{\phi}$	0.5771	0.1875	0.2498	0.0625
	(3.64e-5)	(4.76e-6)	(1.52e-8)	(2.08e-9)
β	0.2500	1.0000	1.0000	0.2500
	(0.0002)	(1.64e-9)	(1.05e-9)	(1.18e-8)
σ_e^2	4.39e-7	1.53e-7	3.05e-7	5.08e-5
	(5.78e-17)	(7.59e-11)	(2.96e-9)	(1.36e-10)

FIGURE 2. Implied volatility of the instantaneous interest rate

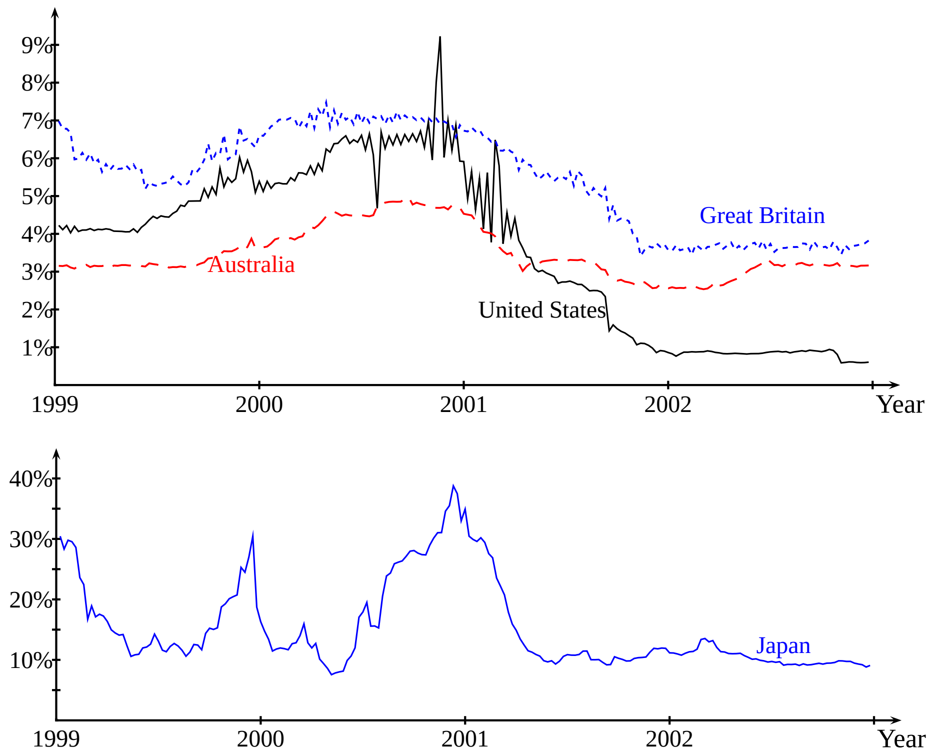
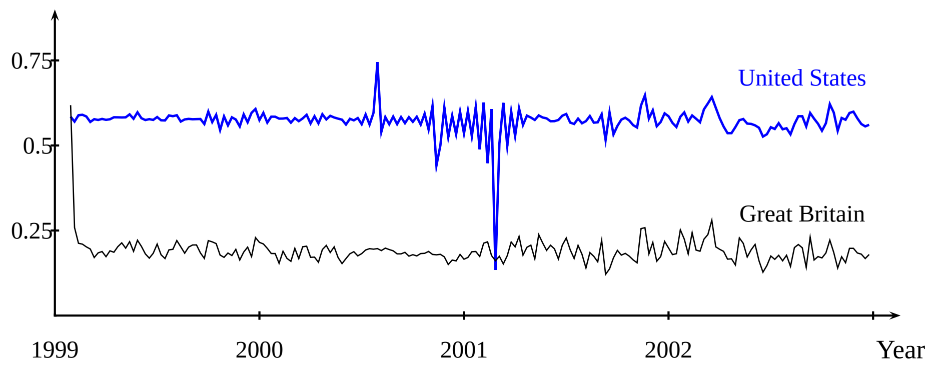
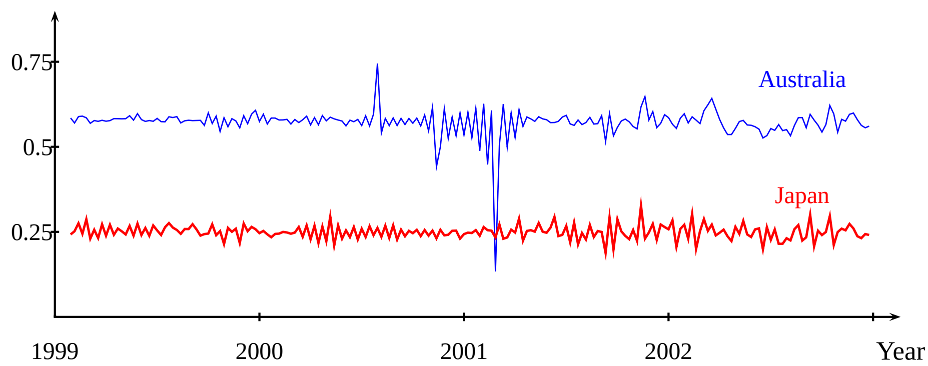


FIGURE 3. Implied market price of interest rate risk



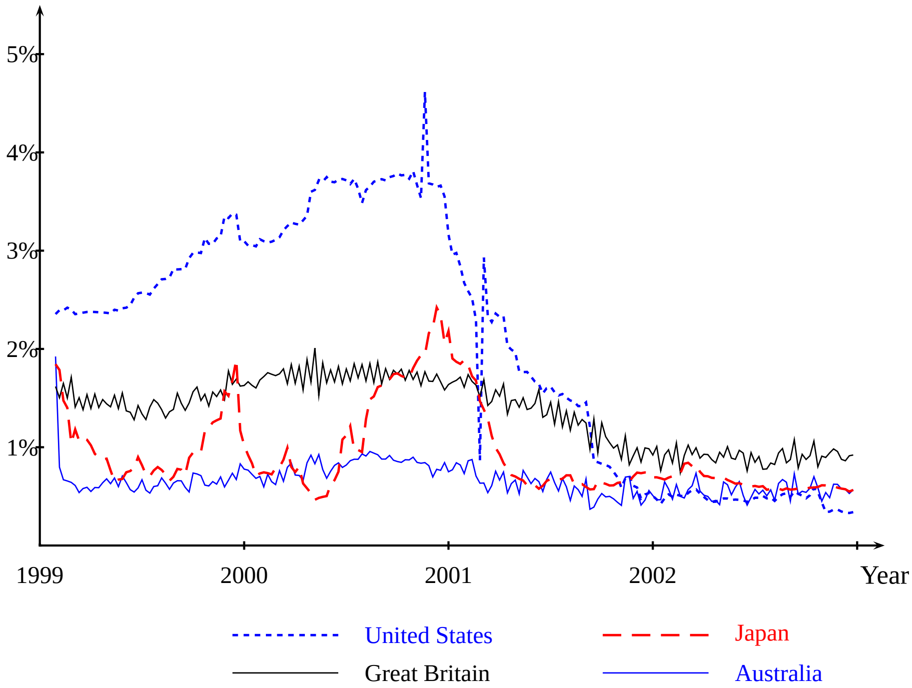


days in the U.S. and the Australian markets and 3.5 days in the Great Britain and the Japanese markets for the market price of risk to half its level.

The discount on the instantaneous short rate (which will be transformed into a premium on bond price) can be determined by scaling the volatility level by this market price of risk. It can be seen from Figure 4 that the discount in the U.S. market is highest and that in the Australian market is lowest. The discount levels in both the Australian and the Great Britain market are stable, whereas that in the Japanese market fluctuates strongly.

In all of the markets the model provides a reasonable fit to data. Table 2 reports the absolute difference between the predicted values for LIBOR rates based on fitted state variables and the observed rates. Overall, the absolute error in the U.S market is 6.7 basis points, and 1.6-3 basis points in the other three markets. The prediction error is similar across the terms of the rates, except for the longest and shortest terms where the errors are slightly higher. The higher errors observed in the U.S market are concentrated on the two quarters from the end of 2000 to the beginning of 2001, where interest rate movement changes its direction from an increase trend to a sharp decline.

FIGURE 4. Implied discount for the instantaneous short rate



The error patterns can be seen in Figure 5, where the errors for 4-month LIBOR rate prediction have been illustrated (the term is chosen so as its average error matches that across maturities).

5. CONCLUSION

The paper has implemented a local linearization filter of Jimenez and Ozaki (2002, 2003) to estimate (via the maximum likelihood method) a model for the LIBOR rates traded in the U.S., Great Britain, Australia and Japan. The model allows a nonlinear dependence of volatility on the level of the interest rate. The estimation predicts a different power dependence than a square-root model usually used in the literature. Interesting behaviours of interest rate and the market price of its risk are also revealed.

FIGURE 5. Prediction error for LIBOR rates

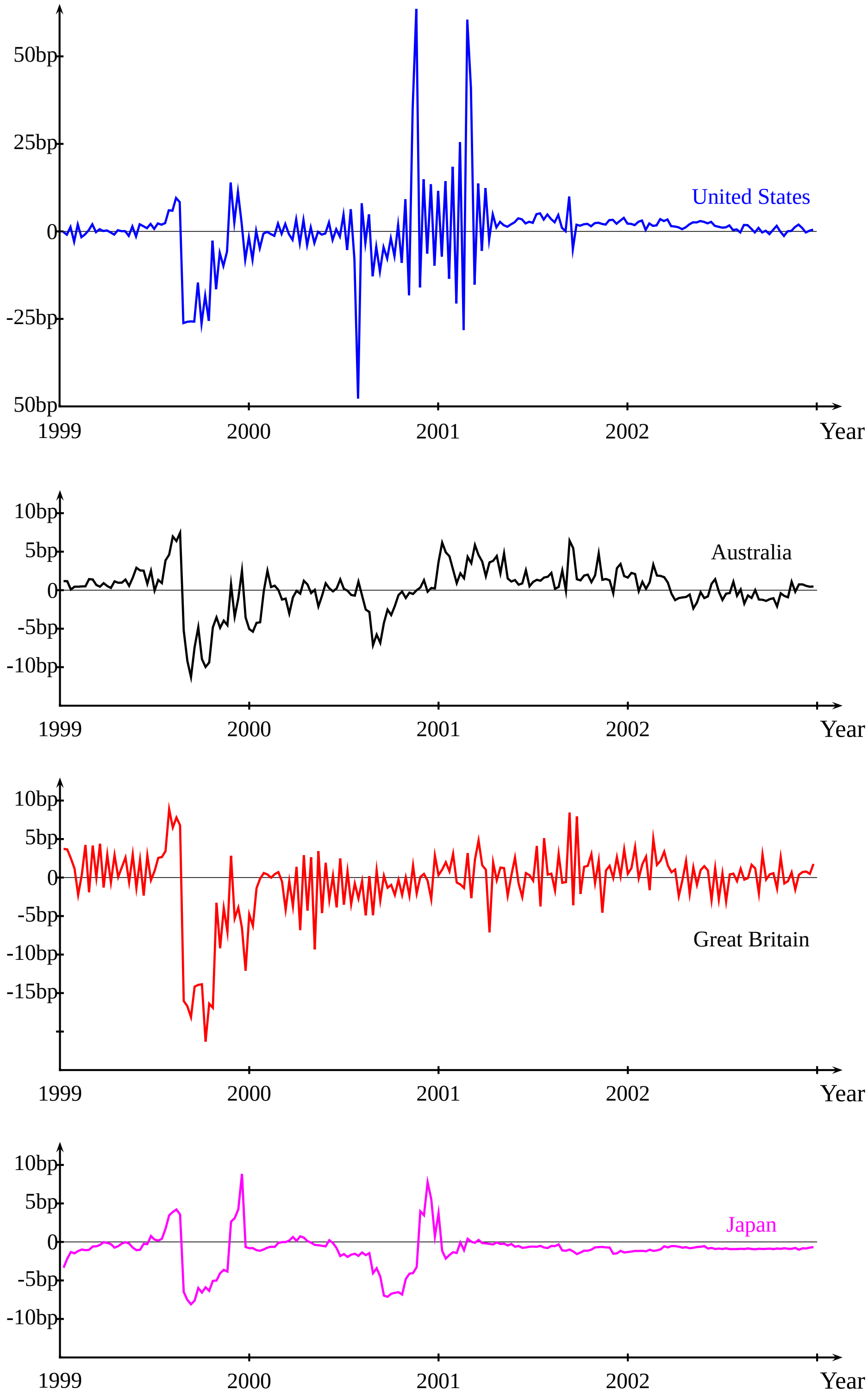


TABLE 2. Absolute difference between the fitted and the observed LIBOR rates

Term	U.S.		Australia		Great Britain		Japan	
(months)	Avr	Stdev	Avr	Stdev	Avr	Stdev	Avr	Stdev
1	0.00132	0.00195	0.00049	0.00042	0.00059	0.00059	0.00031	0.00052
2	0.00089	0.00152	0.00023	0.00020	0.00036	0.00036	0.00021	0.00025
3	0.00069	0.00125	0.00016	0.00022	0.00031	0.00037	0.00019	0.00022
4	0.00058	0.00093	0.00021	0.00021	0.00029	0.00036	0.00016	0.00019
5	0.00052	0.00066	0.00026	0.00021	0.00027	0.00030	0.00013	0.00016
6	0.00044	0.00037	0.00030	0.00022	0.00027	0.00023	0.00011	0.00013
7	0.00036	0.00024	0.00024	0.00017	0.00020	0.00015	0.00007	0.00010
8	0.00038	0.00045	0.00018	0.00013	0.00016	0.00012	0.00006	0.00007
9	0.00045	0.00080	0.00012	0.00008	0.00016	0.00013	0.00009	0.00007
10	0.00056	0.00116	0.00008	0.00006	0.00022	0.00017	0.00013	0.00008
11	0.00078	0.00150	0.00021	0.00015	0.00033	0.00024	0.00018	0.00012
12	0.00105	0.00183	0.00036	0.00025	0.00044	0.00032	0.00024	0.00015
All	0.00067	0.00122	0.00024	0.00024	0.00030	0.00033	0.00016	0.00022

APPENDIX A. 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

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

and the discrete observation equation[8]

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

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 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 random vector i.i.d.

Define $\hat{\mathbf{x}}_{t|\rho} = \mathbb{E}\,(\mathbf{x}(t)|Z_\rho)$ 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 $\rho \le t$, where $Z_\rho = \{\mathbf{z}_{t_j} : t_j \le \rho\}$.

[8]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 view of most finance applications, we chose to omit the term $\boldsymbol{\xi}$.

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 A.1. *(Jimenez and Ozaki (2002)) The optimal (minimum variance) linear filter for the linear model (A.1)- (A.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*

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

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

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

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

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

where

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

is the filter gain. The prediction $\hat{\mathbf{x}}_{t|\rho}$ and $\mathbf{P}_{t|\rho}$ are accomplished, respectively, via expressions (A.3) and (A.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 these 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

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

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.

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.

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.

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.

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.

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.

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.