

A Type of HJM Based Affine Model: Theory and Empirical Evidence *

HAITAO LI AND XIAOXIA YE[†]

*First draft: January 26, 2009, current version: May 18, 2010.

[†]Li is at Stephen M. Ross School of Business, University of Michigan, USA. Ye is at Wang Yanan Institute for Studies in Economics, Xiamen University, China.

Abstract

In this paper a type of Heath, Jarrow and Morton (1992) (HJM) based affine model is derived theoretically. This type of affine model is obtained by applying Linear Realization Theory to construct *Finite Dimensional Realizations* (FDRs) of the Gaussian HJM model. The algorithms of constructing Standard Observable Canonical Realization and Jordan Canonical Realization are introduced sequentially. And it is shown that the commonly adopted FDR is actually Jordan Canonical Realization. The empirical results show that a two-factor model of this type provides great fit to the term structure of interest rates data. The resulting state variables have clear economic interpretations. And it is found that the short end of the term structure can be precisely considered as a “medium-run factor” which uniformly shifts the yield curve. This finding has an important implication for bond portfolios management, and also helps us better understand the interactions between macro-economy and term structure dynamics.

Keywords: Affine Term Structure Model; HJM; Finite Dimensional Realization; Linear Realization Theory; State Space Framework; Macro-economy

JEL classification: C61; E43; E44; G12

1 Introduction

Heath, Jarrow and Morton (1992) (hereafter HJM) term structure of interest rate framework has been proved to be one of the most general and flexible settings for describing the interest rate dynamics, as it is able to generate a wide range of forward rate dynamics and incorporate any prevailing market conditions with internal consistency. Furthermore, under the HJM framework, as less as two ingredients, the volatility function and the initial forward rate curve, are all needs to fully specify the model. Therefore, it is desired that one could link the HJM model with the state space framework (see, e.g., Lemke, 2005), under which models can be easily estimated. There have been series of papers focus on the Markovian property and Finite Dimensional Realizations (hereafter *FDRs*) of HJM, such as Carverhill (1994), Ritchken and Sankarasubramanian (1995), Bliss et al. (1996), Bhar and Chiarella (1996), Björk and Gombani (1999), Björk and Svensson (2001), Björk and Landén (2002), etc. Thanks to the literature, one can analyze the HJM model under the state space framework.

In the literature, there are papers (see Bhar and Chiarella, 1996; Björk and Gombani, 1999, etc.) deal with the hump shape volatility term structure which has been widely observed in many empirical works (see, e.g., Brace and Musiela, 1994; Ritchken and Chuang, 2000; Mercurio and Moraleda, 2001, 2000; Jarrow et al., 2007). HJM with the following time invariant hump shape volatility function

$$\sigma(t, T) = \sigma(T - t) = \left(a_0 + a_1(T - t) + a_2(T - t)^2 + \cdots + a_n(T - t)^n \right) e^{-k(T-t)} \quad (1)$$

where T is the maturity date, t is the current time, and $k > 0$, is guaranteed to admit an *FDR*, because (1) satisfies the sufficient condition for HJM to be represented in an *FDR* (refer to Bhar and Chiarella, 1996, Proposition 1 or Björk and Gombani, 1999, Proposition 3.1 for more details). This paper focuses on *FDRs* associated with the

volatility function of (1). Since given (1) the HJM model is Gaussian, in an *FDR* the forward rate is an affine function of state variables, in other words, the resulting system of an *FDR* can be considered as an affine term structure model. As we'll see in Section 7, even with the volatility function as simple as (1), the resulting affine models are able to well capture the dynamics of the term structure of interest rates.

In Björk and Gombani (1999) Linear System Theory was firstly introduced to solve the *FDR* issues of the HJM model. Although Björk and Gombani (1999) actually construct a concrete *FDR* of HJM in an example, they talk much less on the construction of *FDRs* than the existence of *FDRs*. While Bhar and Chiarella (1996) do provide a more general discussion on how to actually construct an *FDR* of HJM with (1) type of volatility function, they employ a totally different method and reach a different result in appearance. Therefore, one wonders given one HJM whether there exist two or more different *FDRs*, and whether all these realizations are equivalent under certain unifying theoretical framework. In order to answer these questions, this paper provides a method, which further extends Björk and Gombani (1999), to construct *FDRs* of HJM by applying Linear Realization Theory. And it is found that there do exist non-unique *FDRs* that are equivalent to each other given one HJM model, and the canonical forms (Standard Observable Canonical and Jordan Canonical forms) can be easily achieved. Actually, the result in Bhar and Chiarella (1996, Prop. 1) is the Jordan Canonical *FDR* which can be reached by applying the block diagonalization algorithm to a system matrix of state variables.

In the empirical analysis, it is shown that, in terms of the latitude of pricing errors, a two-factor model of this type with only 7 model parameters provides much better fit to the term structure data than a traditional two-factor affine model with 13 model parameters does, and the fit is comparable to that of a traditional three-factor affine model with 21 model parameters (see, e.g., De Jong, 2000). Although there

is no clear economic interpretation for the second state variable¹ of Jordan Canonical Realization, the economic interpretations for the first and third state variables of Jordan Canonical Realization, as well as the three state variables of Standard Observable Canonical Realization are apparent: a linear combination of the first and third states of Jordan Canonical Realization, and the first state of Standard Observable Canonical Realization are almost completely linked to the short end of the term structure (3-month interest rate), and a linear combination of the second and third states of Standard Observable Canonical Realization is highly related to the slope (10-year interest rate minus 3-month interest rate) of the term structure. Generally speaking, the traditional two-factor affine model is only able to capture the dynamics of the slope and level (slope plus short end) of the term structure and barely able to explain the curvature (two times the 1-year zero yield, minus the sum of the 3-month and 10-year zero yields) dynamics. In contrast to this, the two-factor model considered here can not only well capture the slope and level dynamics, but also explain over 80% variation of the curvature.

And using Standard Observable Canonical Realization, it is found that the short end of the term structure can be precisely regarded as a “medium-run factor” which uniformly shifts the yield curve up to 10-year maturity. This finding is of the important implication for the short- and medium-term bond portfolios management. And we all know that the short end of the yield curve is highly sensitive to the macro-economic environment, so this finding also helps us better understand the interactions between the macro-economy and term structure of interest rates dynamics.

The structure of the paper is as follows. First of all, the HJM framework is summarized in Section 2. The general Linear Realization Theory for HJM is briefly re-

¹Although there are only two sources of randomness (stochastic factors), there are three or four relevant state variables in the resulting affine model. This becomes clear as we move on to the succeeding sections

viewed in Section 3; and two representation algorithms (Standard Observable Canonical and Jordan Canonical) of constructing *FDRs* of the HJM model with (1) type volatility function are introduced in Section 4. Bhar and Chiarella (1996)'s method of constructing an *FDR* is presented in Section 5. Then it is shown in Section 6, via a concrete example, that the *FDR* constructed in Bhar and Chiarella (1996) is actually the Jordan Canonical Realization. Based on the same concrete model, Section 7 discusses the empirical results. Finally, Section 8 concludes the paper.

2 The HJM Framework

For completeness, the general HJM model is presented first. Then, in the next section the HJM with the time invariant hump shape volatility function is discussed.

The instantaneous forward rate at time t for all dates $T > t$, $f(t, T)$, is defined by $f(t, T) = -\partial \ln P(t, T) / \partial T$, which is the rate that can be contracted at time t for instantaneous riskless borrowing or lending at time T . From the knowledge of the instantaneous forward rates for all maturities between time t and time T , the price at time t of a zero-coupon bond with maturity T can be obtained by

$$P(t, T) = \exp \left\{ - \int_t^T f(t, s) \, ds \right\},$$

and the corresponding zero yield is

$$y(t, T) = -\frac{\ln P(t, T)}{T - t} = \frac{\int_t^T f(t, s) \, ds}{T - t}.$$

The spot interest rate at time t , r_t , is the instantaneous forward rate at time t for date t , $r_t = f(t, t)$.

Following Heath et al. (1992), the dynamics of the term structure of interest rates

are described by a family of stochastic processes representing forward rate movements. In integral form, the dynamic of the forward rate is given by

$$f(t, T) = f(0, T) + \int_0^t \mu(s, T) \mathbf{d}s + \int_0^t \sigma(s, T) \mathbf{d}w_s$$

or, in SDE form,

$$\mathbf{d}f(t, T) = \mu(t, T) \mathbf{d}t + \sigma(t, T) \mathbf{d}w_t$$

where $\mu(t, T)$ is the instantaneous drift and $\sigma(t, T)_{1 \times N}$ is the instantaneous volatility at time t of the forward rate with maturity date T , $w_{N \times 1}$ is an N -dimensional P measure Wiener process. The processes μ and σ are adapted processes in \mathcal{R} and \mathcal{R}^N such that the forward rate processes are well defined. Heath et al. (1992) show that arbitrage-free instantaneous forward rate processes must verify a constraint on the drift,

$$\mu(t, T) = \sigma(t, T) \left(\int_t^T \sigma(t, s) \mathbf{d}s - \lambda(t)^\top \right)^\top,$$

where $\lambda(t)_{N \times 1}$ is a vector of market prices of risks. Note that the volatility function $\sigma(t, T)$ and market price of risks vector $\lambda(t)$ completely determine the drift of the forward rate.

The dynamics of the spot rate and the zero-coupon bond price are given by

$$\mathbf{d}r_t = \left. \frac{\partial f(t, T)}{\partial T} \right|_{T=t} \mathbf{d}t + \sigma(t, t) \mathbf{d}w_t, \quad (2)$$

$$\mathbf{d}P(t, T) = P(t, T) [r_t + \Omega(t, T)] \mathbf{d}t - P(t, T) \left[\int_t^T \sigma(t, s) \mathbf{d}s \right] \mathbf{d}w_t \quad (3)$$

where

$$\Omega(t, T) \equiv - \int_t^T \mu(t, s) \mathbf{d}s + \frac{1}{2} \left\| \int_t^T \sigma(t, s) \mathbf{d}s \right\|^2.$$

In general, neither the spot rate nor the bond price process is Markovian, because the drift term in (2), $\left. \frac{\partial f(t, T)}{\partial T} \right|_{T=t}$, involves integration over the past. Consequently, the general HJM model does not readily lend itself to practical implementations. However, as shown below, if the volatility function is restricted to certain form, the whole HJM framework can be transformed into a Markovian state space.

3 Linear System Theory and FDRs of HJM

Now the volatility function is restricted to the form of (1), then at least one *FDR* is guaranteed to exist for HJM as discussed in Section 1. Here the main results given in Björk and Gombani (1999) are briefly reviewed without dealing with too much technical details. Readers could refer to original paper for more details.

First, rewrite the standard HJM drift condition in the Musiela parameterization (see Brace and Musiela, 1994):

$$\begin{aligned} \mathbf{d}r(t, x) &= \mu_r(t, x) \mathbf{d}t + \sigma(x) \mathbf{d}W_t & (4) \\ \mu_r(t, x) &= \frac{\partial}{\partial x} r(t, x) - \sigma(x) \left(\int_0^x (\lambda(t)^\top - \sigma(s)) \mathbf{d}s \right)^\top \\ W_t &= [[w_i]_{i=1}^m]^\top \\ \sigma(x) &= [\sigma_i(x)]_{i=1}^m \\ \lambda(t) &= [[\lambda_i(t)]_{i=1}^m]^\top, \end{aligned}$$

where $x > 0$ is the time to maturity. Denote by \mathbf{F} the linear operator $\frac{\partial}{\partial x}$, using

following property of \mathbf{F}

$$\left[e^{\mathbf{F}t} g \right] (x) = g(t+x) \quad g \in C[0, \infty),$$

the solution of (4) can be obtained as

$$\begin{aligned} r(t, x) &= \delta(t, x) + r_0(t, x) \\ \delta(t, x) &= r(0, t+x) \\ &\quad + \int_0^t \left[\sigma(x+t-s) \left(\int_0^{x+t-s} (\sigma(v) - \lambda(t)^\top) \mathbf{d}v \right)^\top \right] \mathbf{d}s \\ r_0(t, x) &= \int_0^t \sigma(x+t-s) \mathbf{d}W_s \end{aligned} \quad (5)$$

$$\mathbf{d}r_0(t, x) = \mathbf{F}r_0(t, x) \mathbf{d}t + \sigma(x) \mathbf{d}W_t, \quad r_0(0, x) = 0 \quad (6)$$

Following Björk and Gombani (1999), we have the definition of the *FDR*,

Definition 1 A triple $\{\mathbf{A}, \mathbf{B}, \mathbf{C}(x)\}$, where \mathbf{A} is an $n \times n$ -matrix, \mathbf{B} is an $n \times m$ -matrix and $\mathbf{C}(x)$ is an n -dimensional row-vector function, is called an n -dimensional realization of the system $r_0(t, x)$ if $r_0(t, x)$ has the representation

$$r_0(t, x) = \mathbf{C}(x) Z_t \quad (7)$$

$$\mathbf{d}Z_t = \mathbf{A}Z_t \mathbf{d}t + \mathbf{B} \mathbf{d}W_t. \quad (8)$$

Then according to Lemma 3.1 and Corollary 3.1 in Björk and Gombani (1999), it is known that, in order to make (7)-(8) a realization of (5), (7)-(8) and (5)-(6) must have the same transfer function.

The transfer function $H(s, x)$ of (7)-(8) is $\mathbf{C}(x) (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}$, and the transfer function $G(s, x)$ of (5)-(6), by Lemma 3.2 in Björk and Gombani (1999), is $\mathcal{L}[\sigma(x)](s)$ where $\mathcal{L}[\cdot]$ is the Laplace transform operator. Then we need to determine \mathbf{A} , \mathbf{B} , and

$\mathbf{C}(x)$ from the following equation

$$\mathcal{L}[\sigma(x)](s) = \mathbf{C}(x)(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B}.$$

Lemma 3.3 in Björk and Gombani (1999) tells us that it is enough to consider the transfer function $G(s, 0) = \mathcal{L}[\sigma(x=0)](s)$, and then \mathbf{A} , \mathbf{B} , and $\mathbf{C}(x)$ can be determined from follows,

$$\begin{aligned} \mathcal{L}[\sigma(x=0)](s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} \\ \mathbf{C}(x) &= \mathbf{C} \times \exp(\mathbf{A}x). \end{aligned}$$

Computing \mathbf{A} , \mathbf{B} , and \mathbf{C} is no other than the realization construction problem which will be detailedly discussed in the next section.

4 Realization Algorithms

The volatility function is rewritten as

$$\left[\sigma_i(x) = \left(a_{i,0} + a_{i,1}x + a_{i,2}x^2 + \cdots + a_{i,n_i}x^{n_i} \right) e^{-k_i x} \right]_{i=1}^m, \quad (9)$$

Then the transfer function $G(s, 0) = \mathcal{L} [[\sigma_i(x=0)]_{i=1}^m](s)$ is given by

$$\begin{aligned}
 & \underbrace{\left[\sum_{j=0}^{n_i} j! a_{ij} (k_i + s)^{-(j+1)} \right]_{i=1}^m}_{1 \times m} \tag{10} \\
 & = \underbrace{\left[\frac{1}{n_i! (k_i + s)^{n_i+1}}, \frac{1}{(n_i - 1)! (k_i + s)^{n_i}}, \dots, \frac{1}{k_i + s} \right]_{i=1}^m}_{1 \times (m + \sum_{i=1}^m n_i)} \cdot \underbrace{\begin{bmatrix} a_{10} & 0 & 0 \\ \vdots & \vdots & \\ a_{1n_1} & 0 & \\ 0 & a_{20} & \vdots \\ & \vdots & \vdots \\ & a_{2n_2} & 0 \\ \vdots & 0 & a_{m0} \\ & \vdots & \vdots \\ 0 & 0 & a_{mn_m} \end{bmatrix}}_{(m + \sum_{i=1}^m n_i) \times m}.
 \end{aligned}$$

The realization algorithms are then applied to

$$\left[\frac{1}{n_i! (k_i + s)^{n_i+1}}, \frac{1}{(n_i - 1)! (k_i + s)^{n_i}}, \dots, \frac{1}{k_i + s} \right]_{i=1}^m. \tag{11}$$

As far as the realization algorithms are concerned, Standard Observable Canonical Realization algorithm is a natural choice for the row vector transfer function (11).

4.1 Standard Observable Canonical Realization

Details of Standard Observable Canonical Realization algorithm can be found in Brockett (1970, page 107). First, expand (11) about $|s| = \infty$ to get

$$\begin{aligned} & \left[\frac{1}{0! (k_1 + s)}, \frac{1}{1! (k_1 + s)^2}, \frac{1}{2! (k_1 + s)^3}, \dots, \frac{1}{n_i! (k_i + s)^{n_i+1}} \right]_{i=1}^m \\ &= \sum_{\alpha=0}^{\infty} [\mathbf{L}_{\alpha 1}, \mathbf{L}_{\alpha 2}, \dots, \mathbf{L}_{\alpha n}] s^{-(\alpha+1)}. \end{aligned}$$

$\mathbf{L}_{\alpha i}$, $i = 1, 2, \dots, n$, $\alpha = 0, 1, \dots$, can be easily computed for any specific examples by applying the Taylor expansion.

Secondly, consider $p(s)$ which is the monic least common multiple of all $n_i! (k_i + s)^{n_i+1}$ for $i = 1, 2, \dots, m$, we have

$$\begin{aligned} p(s) &= \prod_{i=1}^m n_i! (k_i + s)^{n_i+1} \\ &= s^n + p_{n-1} s^{n-1} + \dots + p_1 s + p_0 \end{aligned}$$

where $n = m + \sum_{i=1}^m n_i$. As $\mathbf{L}_{\alpha i}$ above, p_k , $k = 0, 2, \dots, n-1$, can also be easily computed for any specific examples by expanding $\prod_{i=1}^m n_i! (k_i + s)^{n_i+1}$.

Then

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -p_0 & -p_1 & -p_2 & \cdots & -p_{n-1} \end{bmatrix}_{n \times n}$$

$$\dot{\mathbf{B}} = \begin{bmatrix} \mathbf{L}_{01} & \mathbf{L}_{02} & \cdots & \mathbf{L}_{0m} \\ \mathbf{L}_{11} & \mathbf{L}_{12} & \cdots & \mathbf{L}_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{L}_{n-2,1} & \mathbf{L}_{n-2,2} & \cdots & \mathbf{L}_{n-2,m} \\ \mathbf{L}_{n-1,1} & \mathbf{L}_{n-1,2} & \cdots & \mathbf{L}_{n-1,m} \end{bmatrix}_{n \times n}$$

$$\mathbf{B} = \dot{\mathbf{B}}_{n \times n} \times \mathbf{\Omega}_{n \times m}$$

$$\mathbf{C}(x) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}_{1 \times n} \times \exp(\mathbf{A}x).$$

where

$$\mathbf{\Omega}_{n \times m} \equiv \begin{bmatrix} a_{10} & 0 & 0 \\ \vdots & \vdots & \\ a_{1n_1} & 0 & \\ 0 & a_{20} & \vdots \\ & \vdots & \vdots \\ & a_{2n_2} & 0 \\ \vdots & 0 & a_{m0} \\ & \vdots & \vdots \\ 0 & 0 & a_{mn_m} \end{bmatrix}_{n \times m}$$

Thanks to the structure of the transfer function and Standard Observable Canonical Realization algorithm, we have the following nice result:

Proposition 2 *Given the volatility function as*

$$\left[\sigma_i(x) = \left(a_{i,0} + a_{i,1}x + a_{i,2}x^2 + \dots + a_{i,n_i}x^{n_i} \right) e^{-k_i x} \right]_{i=1}^m,$$

triple $\{\mathbf{A}, \mathbf{B}, \mathbf{C}(x)\}$ is a minimal realization.

Proof. The McMillan degree (see Chen, 1999) of the transfer function is given by

$$\mu = \dim \left[\text{span} \left\{ \left[e^{-k_i x}, x e^{-k_i x}, \dots, x^{n_i} e^{-k_i x} \right]_{i=1}^m \right\} \right] = m + \sum_{i=1}^m n_i = n,$$

and the dimension of \mathbf{A} is $n \times n$, then the result follows. ■

Finally, the Markovian state variables and $r_0(t, x)$ are given by

$$\begin{aligned} r_0(t, x) &= \mathbf{C}(x) Z_t \\ \mathbf{d}Z_t &= \mathbf{A}Z_t \mathbf{d}t + \mathbf{B} \mathbf{d}W_t \end{aligned}$$

where $W_t = [[w_i]_{i=1}^m]^\top$.

4.2 Jordan Canonical Realization

Given the Standard Observable Canonical Realization triple $\{\mathbf{A}, \mathbf{B}, \mathbf{C}(x)\}$, Jordan Canonical Realization can be constructed as follows:

First, apply the Jordan form algorithm (see, e.g., Chen, 1999) to matrix \mathbf{A} to obtain

a block-diagonal and triangular-form representation, i.e.

$$\mathbf{A} = \mathbf{P}^{-1}\tilde{\mathbf{A}}\mathbf{P}$$

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{\mathbf{A}}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\mathbf{A}}_2 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & \cdots & \tilde{\mathbf{A}}_m \end{bmatrix}_{n \times n}, \tilde{\mathbf{A}}_i = \begin{bmatrix} -k_i & 0 & \cdots & 0 \\ 1 & -k_i & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & n_i & -k_i \end{bmatrix}_{(n_i+1) \times (n_i+1)}$$

where \mathbf{P} is a unique nonsingular matrix, and $\tilde{\mathbf{A}}$ is a block-diagonal matrix.

Secondly, by the similarity transformation (see, e.g., Schutter, 2000), we obtain the Jordan Canonical Realization triple $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}(x)\}$ as

$$\begin{aligned} \tilde{\mathbf{A}} &= \mathbf{P}\mathbf{A}\mathbf{P}^{-1} \\ \tilde{\mathbf{B}} &= \mathbf{P}\mathbf{B}\mathbf{\Omega} \\ \tilde{\mathbf{C}}(x) &= \tilde{\mathbf{C}} \exp(\tilde{\mathbf{A}}x) \\ &= \mathbf{C}\mathbf{P}^{-1} \exp(\mathbf{P}\mathbf{A}\mathbf{P}^{-1}x) \\ &= \mathbf{C} \exp(\mathbf{A}x) \mathbf{P}^{-1} \\ &= \mathbf{C}(x) \mathbf{P}^{-1} \end{aligned}$$

To see how matrix \mathbf{P} could look like, a specific example is presented in the next section. Note that the triple $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}(x)\}$ is also a minimal realization. Therefore Standard Observable Canonical Realization and Jordan Canonical Realization are equivalent in terms of the controllability and observability (see Schutter, 2000).

4.3 Example Illustration

Here two algorithms are demonstrated via a concrete example. Assume a two-factor model with the volatility function as

$$\sigma(x) = [(a_{10} + a_{11}x) \exp(-k_1x), (a_{20} + a_{21}x) \exp(-k_2x)]_{1 \times 2}.$$

The transfer function $G(s, 0) = \mathcal{L}[(a_{10} + a_{11}x) \exp(-k_1x), (a_{20} + a_{21}x) \exp(-k_2x)](s)$ is given by

$$\begin{aligned} & \left[\frac{a_{10}}{k_1 + s} + \frac{a_{11}}{(k_1 + s)^2}, \frac{a_{20}}{k_2 + s} + \frac{a_{21}}{(k_2 + s)^2} \right]_{1 \times 2} \\ &= \left[\frac{1}{k_1 + s}, \frac{1}{(k_1 + s)^2}, \frac{1}{k_2 + s}, \frac{1}{(k_2 + s)^2} \right] \begin{bmatrix} a_{10} & 0 \\ a_{11} & 0 \\ 0 & a_{20} \\ 0 & a_{21} \end{bmatrix}, \end{aligned}$$

the realization algorithms are applied to

$$\left[\frac{1}{k_1 + s}, \frac{1}{(k_1 + s)^2}, \frac{1}{k_2 + s}, \frac{1}{(k_2 + s)^2} \right]. \quad (12)$$

- Standard Observable Canonical Realization

1. Expand (12) about $|s| = \infty$

$$\begin{aligned} & \left[\frac{1}{k_1 + s}, \frac{1}{(k_1 + s)^2}, \frac{1}{k_2 + s}, \frac{1}{(k_2 + s)^2} \right] \\ &= [1, 0, 1, 0] \frac{1}{s} + [-k_1, 1, -k_2, 1] \frac{1}{s^2} \\ &+ [k_1^2, -2k_1, k_2^2, -2k_2] \frac{1}{s^3} + [-k_1^3, 3k_1^2, -k_2^3, 3k_2^2] \frac{1}{s^4} + \dots; \end{aligned}$$

2. The Monic least common multiple of the denominators is given by

$$\begin{aligned}
p(s) &= (k_1 + s)^2 (k_2 + s)^2 \\
&= k_1^2 k_2^2 + 2(k_1^2 k_2 + k_1 k_2^2) s + (k_1^2 + 4k_1 k_2 + k_2^2) s^2 + 2(k_1 + k_2) s^3 + s^4 \\
&\equiv p_0 + p_1 s + p_2 s^2 + p_3 s^3 + s^4,
\end{aligned}$$

so $n = 4$;

3. \mathbf{A} , \mathbf{B} , and $\mathbf{C}(x)$ can be computed as

$$\begin{aligned}
\mathbf{A} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k_1^2 k_2^2 & -2(k_1^2 k_2 + k_1 k_2^2) & -(k_1^2 + 4k_1 k_2 + k_2^2) & -2(k_1 + k_2) \end{bmatrix}_{4 \times 4} \\
\dot{\mathbf{B}} &= \begin{bmatrix} 1 & 0 & 1 & 0 \\ -k_1 & 1 & -k_2 & 1 \\ k_1^2 & -2k_1 & k_2^2 & -2k_2 \\ -k_1^3 & 3k_1^2 & -k_2^3 & 3k_2^2 \end{bmatrix}_{4 \times 4} \\
\mathbf{B} &= \dot{\mathbf{B}} \mathbf{\Omega}_{4 \times 2} = \begin{bmatrix} a_{10} & a_{20} \\ a_{11} - k_1 a_{10} & a_{21} - k_2 a_{20} \\ k_1^2 a_{10} - 2k_1 a_{11} & k_2^2 a_{20} - 2k_2 a_{21} \\ k_1^2 (3a_{11} - k_1 a_{10}) & k_2^2 (3a_{21} - k_2 a_{20}) \end{bmatrix}_{4 \times 2} \\
\mathbf{C}(x) &= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \\
&\quad \times \exp \left(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k_1^2 k_2^2 & -2(k_1^2 k_2 + k_1 k_2^2) & -(k_1^2 + 4k_1 k_2 + k_2^2) & -2(k_1 + k_2) \end{bmatrix} x \right)
\end{aligned}$$

where

$$\mathbf{\Omega} = \begin{bmatrix} a_{10} & 0 \\ a_{11} & 0 \\ 0 & a_{20} \\ 0 & a_{21} \end{bmatrix}.$$

- Jordan Canonical Realization

1. Applying the Jordan form algorithm to matrix \mathbf{A} gives

$$\mathbf{P}^{-1} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & -k_1 & 1 & -k_2 \\ -2k_1 & k_1^2 & -2k_2 & k_2^2 \\ 3k_1^2 & -k_1^3 & 3k_2^2 & -k_2^3 \end{bmatrix}$$

2. By the similarity transformation, $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$, and $\tilde{\mathbf{C}}(x)$ can be computed as

$$\tilde{\mathbf{A}} = \mathbf{PAP}^{-1} = \begin{bmatrix} -k_1 & 0 & 0 & 0 \\ 1 & -k_1 & 0 & 0 \\ 0 & 0 & -k_2 & 0 \\ 0 & 0 & 1 & -k_2 \end{bmatrix}$$

$$\tilde{\mathbf{B}} = \mathbf{P}\tilde{\mathbf{B}}\mathbf{\Omega}$$

$$\begin{aligned}
&= \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & -k_1 & 1 & -k_2 \\ -2k_1 & k_1^2 & -2k_2 & k_2^2 \\ 3k_1^2 & -k_1^3 & 3k_2^2 & -k_2^3 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 1 & 0 \\ -k_1 & 1 & -k_2 & 1 \\ k_1^2 & -2k_1 & k_2^2 & -2k_2 \\ -k_1^3 & 3k_1^2 & -k_2^3 & 3k_2^2 \end{bmatrix} \begin{bmatrix} a_{10} & 0 \\ a_{11} & 0 \\ 0 & a_{20} \\ 0 & a_{21} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & 0 \\ a_{10} & 0 \\ 0 & a_{21} \\ 0 & a_{20} \end{bmatrix}
\end{aligned}$$

$$\tilde{\mathbf{C}}(x) = \mathbf{C}\mathbf{P}^{-1} \exp(\tilde{\mathbf{A}}x)$$

$$\begin{aligned}
&= \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & -k_1 & 1 & -k_2 \\ -2k_1 & k_1^2 & -2k_2 & k_2^2 \\ 3k_1^2 & -k_1^3 & 3k_2^2 & -k_2^3 \end{bmatrix} \\
&\quad \times \exp \left(\begin{bmatrix} -k_1 & 0 & 0 & 0 \\ 1 & -k_1 & 0 & 0 \\ 0 & 0 & -k_2 & 0 \\ 0 & 0 & 1 & -k_2 \end{bmatrix} x \right) \\
&= \begin{bmatrix} xe^{-xk_1} & e^{-xk_1} & xe^{-xk_2} & e^{-xk_2} \end{bmatrix}.
\end{aligned}$$

5 The General Method for Constructing an FDR

For an m -dimensional HJM of which each volatility has the same functional form as

$$\sigma_i(t, T) = \left(a_{i,0} + a_{i,1}(T-t) + a_{i,2}(T-t)^2 + \cdots + a_{i,n_i}(T-t)^{n_i} \right) e^{-k_i(T-t)} \quad (13)$$

$i = 1, 2, \dots, m$, the forward rate, under P measure, is given by

$$f(t, T) = f(0, T) + \Theta(t, T) + \sum_{i=1}^m \int_0^t \sigma_i(s, T) \mathbf{d}w_i(s), \quad (14)$$

where

$$\Theta(t, x) = \int_0^t \left(\sum_{i=1}^m \sigma_i(s, x) \int_s^x (\sigma_i(s, v) - \lambda_i(t)) \mathbf{d}v \right) \mathbf{d}s; \quad (15)$$

and the log of bond price

$$\begin{aligned} B(t, T) &= - \int_t^T f(t, s) \mathbf{d}s \\ &= B(0, T) - B(0, t) - \int_t^T \Theta(t, s) \mathbf{d}s - \int_t^T \left(\sum_{i=1}^m \int_0^t \sigma_i(s, v) \mathbf{d}w_i(s) \right) \mathbf{d}v. \end{aligned}$$

Note that $\sum_{i=1}^m \int_0^t \sigma_i(s, T) \mathbf{d}w_i(s) = r_0(\cdot, T - \cdot)$ (see (5)).

The main idea of the *general method for constructing an FDR* introduced in Bhar and Chiarella (1996) is summarized in the following proposition. This method has been commonly adopted to construct FDRs of HJM in the literature.

Proposition 3 *If the forward rate volatility function assumes the form*

$$\sigma_i(t, T) = \left(a_{i,0} + a_{i,1}(T-t) + a_{i,2}(T-t)^2 + \cdots + a_{i,n_i}(T-t)^{n_i} \right) e^{-k_i(T-t)},$$

then the forward rate $f(t, T)$ can be expressed as an affine function of Markovian state vari-

ables, i.e.

$$f(t, T) = f(0, T) + \Theta(t, T) + \check{\mathbf{C}}(T - t) X_t \quad (16)$$

$$dX_t = \check{\mathbf{A}}X_t + \check{\mathbf{B}}dW_t \quad (17)$$

where

$$\begin{aligned} X_t^\top &= \underbrace{\left[\left[\int_0^t (t-s)^j e^{-k_i(t-s)} dW_i(s) \right]_{j=0}^{n_i} \right]_{i=1}^m}_{1 \times n}, \\ \check{\mathbf{C}}(x) &= \underbrace{\left[\left[\sum_{l=j}^{n_i} a_{il} \binom{l}{j} x^{l-j} e^{-k_i x} \right]_{j=0}^{n_i} \right]_{i=1}^m}_{1 \times n}, \\ \check{\mathbf{A}} &= \begin{bmatrix} \check{\mathbf{A}}_1 & 0 & \cdots & 0 \\ 0 & \check{\mathbf{A}}_2 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & \cdots & \check{\mathbf{A}}_m \end{bmatrix}_{n \times n}, \quad \check{\mathbf{A}}_i = \begin{bmatrix} -k_i & 0 & \cdots & 0 \\ 1 & -k_i & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & n_i & -k_i \end{bmatrix}_{(n_i+1) \times (n_i+1)} \\ \check{\mathbf{B}} &= \begin{bmatrix} \check{\mathbf{B}}_1 \\ \check{\mathbf{B}}_2 \\ \vdots \\ \check{\mathbf{B}}_m \end{bmatrix}_{n \times m}, \quad \check{\mathbf{B}}_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & & \cdots & \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{(n_1+1) \times m}, \quad \check{\mathbf{B}}_2 = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ & & \cdots & \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{(n_2+1) \times m} \quad \cdots \\ W_t &= [[w_i]_{i=1}^m]^\top \end{aligned} \quad (18)$$

Proof. Considering the last term in RHS of (14), for each i

$$\begin{aligned}
& \int_0^t \sigma_i(s, T) \mathbf{d}w_i(s) \\
&= \sum_{j=0}^{n_i} \int_0^t a_{i,j} (T-s)^j e^{-k_i(T-s)} \mathbf{d}w_i(s) \\
&= e^{-k_i(T-t)} \sum_{j=0}^{n_i} \int_0^t \sum_{l=0}^j \binom{j}{l} (T-t)^{j-l} a_{i,j} (t-s)^l e^{-k_i(t-s)} \mathbf{d}w_i(s) \\
&= e^{-k_i(T-t)} \sum_{j=0}^{n_i} \int_0^t \sum_{l=j}^{n_i} a_{i,l} \binom{l}{j} (T-t)^{l-j} (t-s)^j e^{-k_i(t-s)} \mathbf{d}w_i(s) \\
&= \sum_{j=0}^{n_i} \left[e^{-k_i(T-t)} \sum_{l=j}^{n_i} a_{i,l} \binom{l}{j} (T-t)^{l-j} \right] X_{i,j}(t),
\end{aligned}$$

This completes the proof. ■

Remark 1 Let's consider the situation of one factor and Q measure, i.e. $m = 1$ and $\lambda = \mathbf{0}_{n \times 1}$. The result of Proposition 3 is more general than that introduced in Bhar and Chiarella (1996). Setting $T = t$, (16) is reduced to the short rate

$$r(t) = f(t, t) = f(0, t) + \int_0^t \left(\sigma_1(s, t) \int_{\tau}^t (\sigma_1(s, v)) \mathbf{d}v \right) ds + \sum_{j=0}^{n_1} a_{1j} X_{1,j}(t).$$

And the log of bond price is

$$B(t, T) = - \int_t^T f(t, s) \mathbf{d}s.$$

Considering the SDE of $[B(t, T); r(t); X_{1,0}(t); \dots; X_{1, n_1-1}(t)]_{(n_1+2) \times 1}$, the result of Bhar and Chiarella (1996, Proposition 1) follows immediately. Proposition 3 here, however, can be more readily applied to the state space estimation.

6 The Equivalence between the General FDR and Jordan Canonical Realization

6.1 The General Observation

Note that $\check{C}(x)$ in Equ. (18) can be decomposed as

$$\check{C}(x) = \underbrace{\left[\left[x^{n_i-j} e^{-k_i x} \right]_{j=0}^{n_i} \right]_{i=1}^m}_{1 \times n} \times \mathbf{a}_{n \times n} = \tilde{C}(x) \times \mathbf{a}_{n \times n}$$

where

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{a}_2 & \cdots & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \mathbf{a}_m \end{bmatrix}_{n \times n}, \quad \mathbf{a}_i = \begin{bmatrix} a_{in_i} & 0 & \cdots & 0 \\ a_{i,n_i-1} & a_{i,n_i} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{i0} & a_{i1} & \cdots & a_{in_i} \end{bmatrix}_{(n_i+1) \times (n_i+1)}.$$

Then define a new set of state variables Z_t as

$$Z_t = \mathbf{a} \times X_t,$$

and from (17), we have Z_t 's dynamic as

$$dZ_t = \mathbf{a} \times dX_t = \mathbf{a} \times \check{A}X_t + \mathbf{a} \times \check{B}dW_t.$$

Note that

$$\mathbf{a} \times \check{A} = \check{A} \times \mathbf{a}, \quad \mathbf{a}^{-1} \check{A} \mathbf{a} = \check{A},$$

hence

$$\begin{aligned} \mathbf{d}Z_t &= \check{\mathbf{A}}Z_t + \mathbf{a} \times \check{\mathbf{B}}\mathbf{d}W_t \\ &= \tilde{\mathbf{A}}Z_t + \tilde{\mathbf{B}}\mathbf{d}W_t. \end{aligned}$$

Finally, the system (16)-(17) can be rewritten as

$$\begin{aligned} f(t, T) &= f(0, T) + \Theta(t, T) + \tilde{\mathbf{C}}(T - t)Z_t \\ \mathbf{d}Z_t &= \tilde{\mathbf{A}}Z_t + \tilde{\mathbf{B}}\mathbf{d}W_t \end{aligned}$$

which is exactly the *Jordan Canonical Realization* discussed in Section 4.2. The next subsection makes the equivalence more visualized by re-investigating the example in Section 4.3.

6.2 A Specific Example

Following Section 4.3, let's assume a two-factor model with the volatility function as

$$\sigma(T - t) = [(a_{10} + a_{11}(T - t)) \exp(-k_1(T - t)), (a_{20} + a_{21}(T - t)) \exp(-k_2(T - t))]_{1 \times 2}.$$

Applying Proposition 3 to the volatility function gives us $\{\check{\mathbf{A}}, \check{\mathbf{B}}, \check{\mathbf{C}}(x)\}$

$$\check{\mathbf{A}} = \begin{bmatrix} -k_1 & 0 & 0 & 0 \\ 1 & -k_1 & 0 & 0 \\ 0 & 0 & -k_2 & 0 \\ 0 & 0 & 1 & -k_2 \end{bmatrix}$$

$$\check{\mathbf{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\check{\mathbf{C}}(x) = \left[(a_{10} + a_{11}x) e^{-k_1 x}, a_{11} e^{-k_1 x}, (a_{20} + a_{21}x) e^{-k_2 x}, a_{21} e^{-k_2 x} \right];$$

then \mathbf{a} is given by

$$\mathbf{a} = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{10} & a_{11} & 0 & 0 \\ 0 & 0 & a_{21} & 0 \\ 0 & 0 & a_{20} & a_{21} \end{bmatrix}.$$

Finally, $\{\check{\mathbf{A}}, \check{\mathbf{B}}, \check{\mathbf{C}}(x)\}$ and \mathbf{a} lead us to the *Jordan Canonical Realization* $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}(x)\}$

$$\tilde{\mathbf{A}} = \check{\mathbf{A}}$$

$$\tilde{\mathbf{B}} = \mathbf{a}\check{\mathbf{B}} = \begin{bmatrix} a_{11} & 0 \\ a_{10} & 0 \\ 0 & a_{21} \\ 0 & a_{20} \end{bmatrix}$$

$$\tilde{\mathbf{C}}(x) = \check{\mathbf{C}}(x) \mathbf{a}^{-1} = \left[x e^{-xk_1}, e^{-xk_1}, x e^{-xk_2}, e^{-xk_2} \right].$$

7 Empirical Analysis

7.1 Data

The data used in this paper are the U.S. Treasury constant maturities yield curve rates, downloaded from H.15. The dataset contains weekly (Friday) observations of the yields with maturities of 3, 6-month, 1, 2, 3, 5, 7, and 10-year, and the sample spans 28 years from 08-Jan-1982 to 16-Apr-2010. So totally, there are 1476×8 data points in the dataset. As the Treasury yield curve is considered as the par curve, the par curves are converted to the zero curves for ease of estimating model parameters, using MATLAB® built-in function `pyld2zero`.

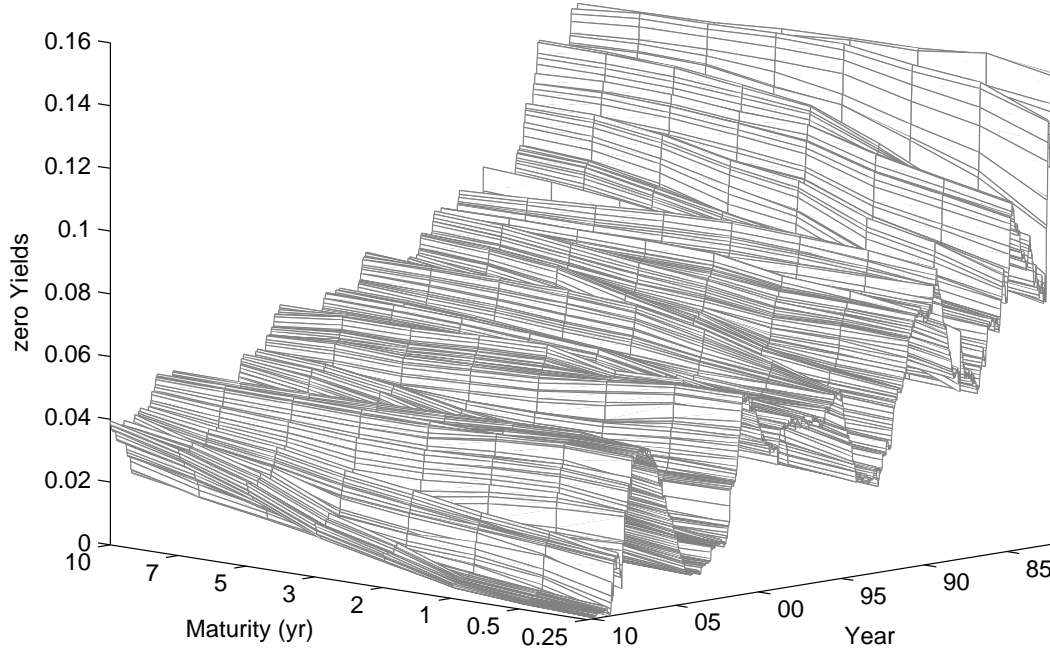
Figure 1 depicts the data, and Table 1 reports the summary statistics on the zero yields. The average interest rates exhibits an upward sloping term structure. The standard deviation of interest rates shows a hump-shaped term structure that peaks around two years. And the interest rates are right-skewed and show excess kurtosis at all the maturities. We could also see, from 1980 to 2010, the level of the interest rate drops down a lot (from around 16% in 1980 to 2% in 2010).

7.2 Estimation Method

Here the concrete model introduced in Section 4.3 is applied to the data. The model is time-inhomogeneous, as the dynamics of the forward rate curve depend on $\delta(t, x)$ (see (5)) which consists of initial term structure $r(0, t + x)$ and volatility component $\Theta(t, x)$ (see (15)). For the purpose of estimation, a constant vector of market prices of risks is assumed, i.e. $\lambda(t) \equiv \lambda$, then the model is reduced to its time-homogeneous counterpart. This can be interpreted as that the model has been evolving from the distant past. Regarding $r(0, t + x)$, a flat initial term structure is assumed (see, e.g., Trolle and Schwartz, 2009; De Jong and Santa-Clara, 1999), specifically, we set

Figure 1: Dynamics of Zero Yields

This figure provides the dynamics of the zero yields with the maturities from 3-month to 10-year. Sample period spans Jan 1982 and Apr 2010.



$r(0, t+x) = \varphi \geq 0$, and φ is estimated as part of the estimation procedure. Regarding the time-inhomogeneous volatility component, $\Theta(t, x)$ is replaced by its limiting counterpart²

$$\begin{aligned}\Theta^*(x) &= \lim_{t \rightarrow \infty} \Theta(t, x) \\ &= \lim_{t \rightarrow \infty} \sum_{i=1}^2 \left[\int_0^t \sigma_i(x+t-s) \left(\int_0^{x+t-s} (\sigma_i(\tau) - \lambda_i) \mathbf{d}\tau \right) \mathbf{d}s \right]\end{aligned}$$

A preliminary estimation shows that estimates for a_{11} and φ are virtually zeros,³

²This setting is different from that of De Jong and Santa-Clara (1999) in which $\Theta(t, x)$ is regarded as an additional state. However, in case of the linear HJM model, $\Theta(t, x)$ is completely deterministic over time, adding an additional state into the state space is more of a hindrance than a help.

³In Standard Observable Canonical Realization, $\hat{a}_{11} = -4.973 \times 10^{-6}$, $\hat{\varphi} = 1.476 \times 10^{-5}$; in Jordan

Table 1: Summary Statistics of Interest Rates

Maturity	Min	1stQ	Medn	3rdQ	Max	Mean	Std.	Skew	Kurt
3m	0.0002	0.0313	0.0512	0.0645	0.1496	0.0504	0.0279	0.3536	3.2159
6m	0.0014	0.0331	0.0530	0.0680	0.1559	0.0530	0.0292	0.4485	3.3583
1y	0.0027	0.0348	0.0546	0.0725	0.1539	0.0551	0.0297	0.5037	3.3221
2y	0.0071	0.0400	0.0575	0.0780	0.1538	0.0593	0.0303	0.5636	3.2322
3y	0.0096	0.0427	0.0588	0.0797	0.1532	0.0616	0.0298	0.6323	3.2216
5y	0.0137	0.0453	0.0613	0.0816	0.1518	0.0652	0.0286	0.7579	3.2381
7y	0.0176	0.0465	0.0633	0.0838	0.1518	0.0681	0.0279	0.8142	3.1857
10y	0.0224	0.0479	0.0638	0.0850	0.1505	0.0698	0.0268	0.8768	3.1629

This table provides the summary statistics of the zero yields for each maturity, which include the minimum, 1st-quantile, median, 3rd-quantile, maximum, mean, standard deviation, skewness, and kurtosis.

therefore the model is refined by dropping a_{11} and φ , and the number of resulting state variables is three, the corresponding $\{\mathbf{A}, \mathbf{B}, \mathbf{C}(x)\}$ and $\{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}(x)\}$ are presented in Appendix A. In order to employ the Kalman filter in conjunction with QMLE technique to estimate the model, the whole model is rewritten in the state space form, and the zero yield is used as the measurement. Specifically, Standard Observable Canonical Realization has the following state space representation,

$$dZ_t = \mathbf{A}Z_t dt + \mathbf{B}dW_t \quad \text{Transition Equation}$$

$$Z_0 = \mathbf{0}_{3 \times 1}$$

$$y_t^x = \varphi + \frac{\int_0^x \Theta^*(s) ds}{x} + \frac{\int_0^x \mathbf{C}(s) ds}{x} Z_t + \epsilon_t \quad \text{Measurement Equation}$$

Canonical Realization, $\hat{a}_{11} = -4.945 \times 10^{-6}$, $\hat{\varphi} = 1.533 \times 10^{-5}$.

or in discrete time form

$$Z_t = \exp(\mathbf{A}h) Z_{t-h} + \exp(\mathbf{A}h) \int_0^h \exp(-\mathbf{A}s) \mathbf{B} dW_s \quad \text{Transition Equation}$$

$$Z_0 = \mathbf{0}_{3 \times 1}$$

$$y_t^x = \varphi + \frac{\int_0^x \Theta^*(s) ds}{x} + \frac{\int_0^x \mathbf{C}(s) ds}{x} Z_t + \epsilon_t \quad \text{Measurement Equation}$$

where y_t^x is the market observation of zero yield vector at time t which is of 8-dimension,⁴ h is the time interval which is one week in this paper, and ϵ_t is the time t measurement error vector which is assumed to be IID normally distributed, i.e.,

$$\epsilon_t \sim \text{IIDN}(\mathbf{0}_{8 \times 1}, \epsilon \mathbf{Eye}_8)$$

\mathbf{Eye}_8 is an 8-dimensional identity matrix, $\mathbf{0}_{. \times 1}$ is a $.$ -dimension vector of zeros, ϵ is a positive free parameter. The state space representation for Jordan Canonical Realization can be written down accordingly, and it is omitted for the sake of brevity.

Under the normal measurement error assumption, the distribution of y_t^x conditional on the information set \mathcal{F}_{t-h} is an 8-dimensional normal distribution with the mean $y_{t|t-h}$ and variance-covariance matrix P_{y_t} . Thus, the transition density of y_t^x can be written as

$$p_{t-h}(y_t^x) = \left[(2\pi)^4 \sqrt{|P_{y_t}|} \right]^{-1} \exp \left[-\frac{1}{2} (y_t^x - y_{t|t-h})^\top P_{y_t}^{-1} (y_t^x - y_{t|t-h}) \right],$$

where $y_{t|t-h}$ and P_{y_t} are outputs from the Kalman filter updating (the detailed procedure is presented in Appendix B). Then the log-likelihood function is given by

⁴There are 8 maturities of zero yields.

$$\ln \mathcal{L} = l \propto - \sum_{i=1}^{1476} \ln |P_{y_{ih}}| - \sum_{i=1}^{1476} \left(y_{ih}^x - y_{ih|(i-1)h} \right)^\top P_{y_{ih}}^{-1} \left(y_{ih}^x - y_{ih|(i-1)h} \right). \quad (19)$$

The model parameters are estimated by maximizing above log-likelihood function.

7.3 Parameter Estimates and Fitting Performance

Table 2 reports the parameter estimates for Standard Observable Canonical Realization (hereafter SOCR) and Jordan Canonical Realization (hereafter JCR). We can see from the table that the two sets of parameter estimates are exactly consistent with each other, this confirms the validity of the theoretical equivalence between SOCR and JCR which has been derived in this paper. Although the standard errors for the estimates of λ_1 and λ_2 are a little larger in JCR than in SOCR, those for k_1 and k_2 are much smaller in JCR than in SOCR. This indicates that k_i s might be easier to identify in JCR than in SOCR. And results not reported show the computation time for JCR is much shorter than that for SOCR. This is natural since, as we can see in Section 4.3, the parametrization of JCR is more concise than that of SOCR. And the concise parametrization of JCR might also explain that k_i s are easier to identify in JCR than in SOCR. Therefore, for a complex model specification, it might be more efficient to employ JCR to estimate parameters.

Figure 2 presents the term structures of the two volatilities calculated based on the parameter estimates in Table 2. We can see that the first volatility increases exponentially as the maturity decreases, this means that the first factor affects short maturity forward rates more than it does long term rates and can be interpreted as a spread between a “short” and “long term factor”; and the second volatility exhibits a hump shape term structure which peaks around 5 years, so the second factor has

Table 2: Estimation Results of two FDRs

FDRs		a_{i0}	a_{i1}	k_i	λ_i	$\sqrt{\epsilon}$	$\ln \mathcal{L}$
SOCR	$i = 1$	0.0106 (0.0001)		0.1146 (0.0013)	-0.4980 (0.0131)	0.0015 (0.0000)	68655
	$i = 2$	0.0019 (0.0006)	-0.0057 (0.0003)	0.2026 (0.0023)	-0.1950 (0.0113)		
JCR	$i = 1$	0.0106 (0.0001)		0.1146 (0.0004)	-0.4980 (0.0139)	0.0015 (0.0000)	68655
	$i = 2$	0.0019 (0.0005)	-0.0057 (0.0003)	0.2026 (0.0006)	-0.1950 (0.0117)		

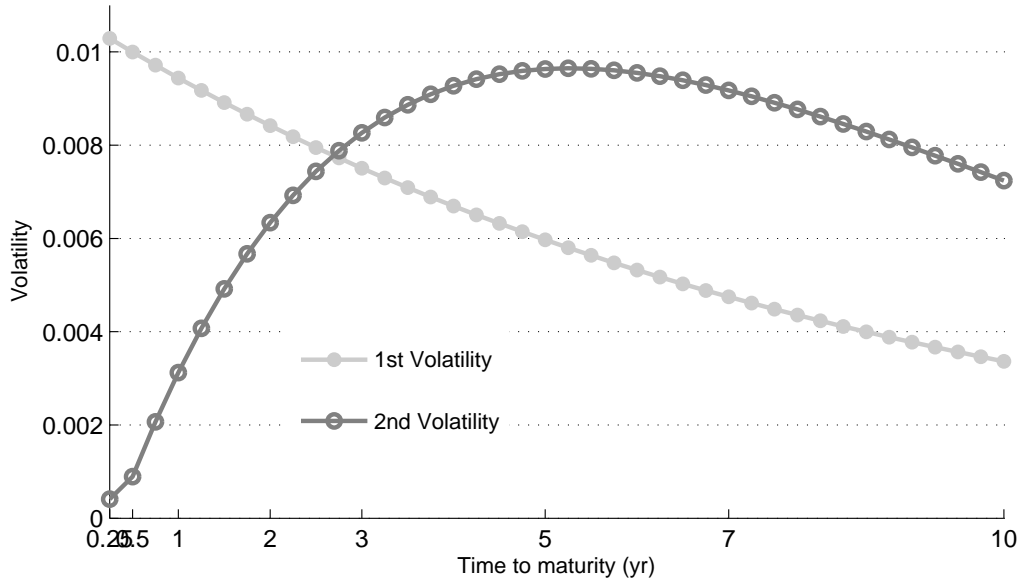
This table reports the QML estimates and standard errors for parameters of the models in SOCR and JCR. The table also reports the square root of measurement error variances $\sqrt{\epsilon}$ s and log-likelihood values $\ln \mathcal{L}$ s.

a curvature effect on the forward rate curve. From Table 2, we find the estimates for market price of risk parameters λ_1 and λ_2 are negative, this indicates that the means of state variables under Q measure are higher than those under P measure, and the risks of these two random factors are positively priced by the market.

Table 3 presents the summary statistics of the pricing errors, and Figure 3 describes the dynamics of the time series of pricing errors. Here the pricing error is defined as the difference between the observed yield and the model implied yield (former minus latter). The pricing errors are also exactly the same for SOCR and JCR, again this confirms the theoretical equivalence between these two realizations. And the statistics show negligible mean and median pricing errors. The average mean pricing errors across different maturities is -0.4 basis point, and the average median is 0.21 basis point. The standard deviations of the pricing errors are all within 20 basis points, and the mean absolute pricing errors are all within 15

Figure 2: Volatilities Term Structure

This figure provides the term structure of the two volatilities computed based on parameter estimates in Table 2. These two volatilities are $|a_{10} \exp(-k_1 x)|$ and $|(a_{20} + a_{21}x) \exp(-k_2 x)|$, where x is the time to maturity. Plots here are for both SOCR and JCR, as the differences between them are negligible.



basis points. It is thus clear that, in terms of latitude of pricing error, the fitting performance documented here is much better than that of the two-factor affine model with 13 parameters, and comparable to that of the three-factor affine model with 21 parameters considered by De Jong (2000).⁵ The last column reports the percentage explained variation, defined as one minus the ratio of the pricing error variance to the variance of the original series, represented in percentage points. The explained variation estimates are over 99.5% for all maturities. Hence, the model captures the term structure of interest rates very well. However, it is also found that the autocorrelations of the pricing errors are relatively high. Figure 3 confirms this observation. This suggests that the model considered here does not perfectly fit the term structure

⁵In De Jong (2000) the average mean pricing errors is 12.64 basis points with average standard deviation of 60.48 basis points for the two-factor, and the average mean pricing errors is -4.58 basis points with average standard deviation of 59.63 basis points for the three-factor.

data yet. And the high autocorrelations might imply that the pricing errors contain the economic information which sheds light on whether the pricing errors are due to the model mis-specification or market inefficiency (see, e.g., Sercu and Xueping, 1997; Turan et al., 2009). We leave this topic for future research.

Table 3: Summary Statistics of Pricing Errors

Maturity	Mean(bp)	Medn(bp)	Std.(bp)	MAE(bp)	Auto	Max(bp)	Min(bp)	VR(%)
3m	-5.80	-4.54	20.50	15.40	0.93	31.82	-146.94	99.46
6m	5.37	5.36	7.78	7.29	0.79	59.81	-29.17	99.93
1y	-0.59	-0.85	12.16	9.20	0.94	68.33	-31.50	99.83
2y	1.55	2.37	15.66	12.57	0.96	78.24	-41.67	99.73
3y	-4.23	-4.46	12.29	10.59	0.94	36.19	-40.90	99.83
5y	-3.34	-1.02	9.69	7.63	0.93	17.37	-44.85	99.89
7y	3.96	4.03	9.13	7.71	0.91	32.89	-45.41	99.89
10y	-0.17	0.80	15.11	12.32	0.96	51.17	-43.71	99.68
Ave.	-0.40	0.21	12.79	10.34	0.92	46.98	-53.02	99.78

This table reports the summary statistics (sample mean, median, standard deviation, mean absolute error, first order autocorrelation, maximum, and minimum) of the pricing errors. The pricing error is defined as the difference between the market observation and the model implied value (former minus latter). The table also reports the variance ratio (VR) at each maturity, which is defined as one minus the ratio of the pricing error variance to the variance of the original series. The results here are for both SOCR and JCR, as there are no distinguishable differences between them.

7.4 State Variables Interpretations

In this subsection, the economic interpretations of the three state variables in both SOCR and JCR are explored.

Figure 3: Time Series of Pricing Errors

This figure depicts the time series of the pricing errors at eight maturities. Plots here are for both SOCR and JCR, as the differences between them are negligible.

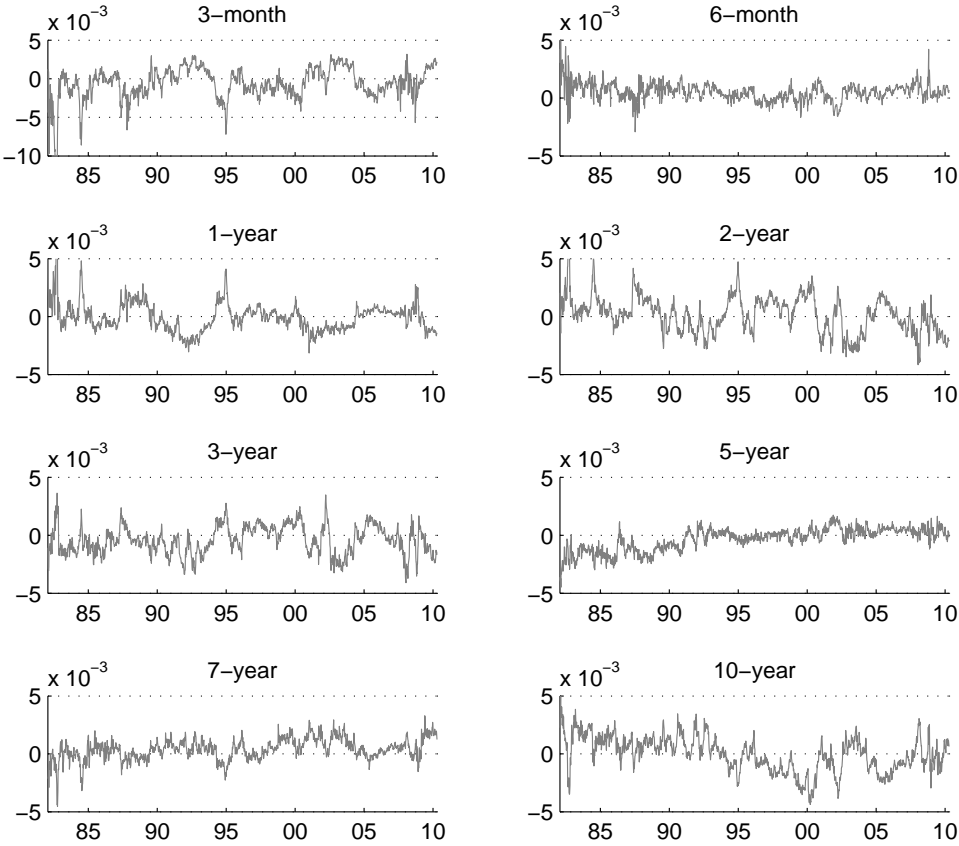


Table 4 reports the summary statistics of the three latent state variables, which are filtered by the Kalman filter based on the model parameter estimates given in Table 2, for both SOCR and JCR. For comparison, the theoretical covariance and correlation matrices are also presented in the table. They can be calculated by inputting the

parameter estimates to the following limitations,

$$\begin{aligned} \text{Cov}_{\text{SOCR}}(Z_t) &= \lim_{h \rightarrow +\infty} \exp(\mathbf{A}h) \left[\int_0^h \exp(-\mathbf{A}s) \mathbf{B}\mathbf{B}^\top \exp(-\mathbf{A}s)^\top \mathbf{d}s \right] \exp(\mathbf{A}h)^\top \\ \text{Cov}_{\text{JCR}}(Z_t) &= \lim_{h \rightarrow +\infty} \exp(\tilde{\mathbf{A}}h) \left[\int_0^h \exp(-\tilde{\mathbf{A}}s) \tilde{\mathbf{B}}\tilde{\mathbf{B}}^\top \exp(-\tilde{\mathbf{A}}s)^\top \mathbf{d}s \right] \exp(\tilde{\mathbf{A}}h)^\top. \end{aligned}$$

Both the means and standard deviations of the first state variable in SOCR are close to those in JCR. However, the second and third state variables are not that similar in SOCR and JCR, although the signs of the means are consistent. In SOCR the first state variable is most volatile, followed by the second and third states; in JCR the first state is also most volatile, and followed by the third and second states. When looking at the theoretical covariance and correlation matrices, in JCR the first state variable is uncorrelated with the other two state variables in theoretical sense. However, as shown in Table 4, the uncorrelation between states in JCR seems too restricted to maintain in the empirical analysis, i.e., the empirical results for covariance and correlation matrices are quite different from the theoretical results. In contrast, in SOCR the three states are correlated with each other in a theoretical sense, then the empirical results in SOCR for covariance and correlation matrices are much more consistent with the theoretical results than the case in JCR. And let's focus on the empirical results of SOCR. It is found that in SOCR the first state variable has more independency relative to the other two states, it is almost uncorrelated with the second state variable, as the correlation between them is only 0.093; and it is mildly inversely related to the third state variable with the correlation of -0.353 . However, the second and third state variables are strongly negatively correlated with each other, as the correlation between them is -0.964 .

Figure 4 describes the dynamics of the three latent state variables for both SOCR and JCR. And Figure 5 presents the factor loading term structures of these three

Table 4: Summary Statistics of Filtered State Variables

		SOCR			JCR		
		1st	2nd	3rd	1st	2nd	3rd
	Mean	-0.1074	0.0139	-0.0020	-0.0992	0.0008	-0.0082
	Std.	0.0287	0.0084	0.0036	0.0320	0.0093	0.0150
CovMatrix (bp,empirical)	1st	0.8218	0.0222	-0.0364	1.0258	0.1054	-0.2149
	2nd	0.0222	0.0700	-0.0290	0.1054	0.0862	0.0119
	3rd	-0.0364	-0.0290	0.0130	-0.2149	0.0119	0.2259
CovMatrix (bp,theoretical)	1st	1.3383	-0.0578	-0.0277	0.4891	0.0000	0.0000
	2nd	-0.0578	0.0518	-0.0191	0.0000	0.0795	0.1703
	3rd	-0.0277	-0.0191	0.0089	0.0000	0.1703	0.8492
CorMatrix (empirical)	1st	1.0000	0.0925	-0.3531	1.0000	0.3546	-0.4465
	2nd	0.0925	1.0000	-0.9637	0.3546	1.0000	0.0851
	3rd	-0.3531	-0.9637	1.0000	-0.4465	0.0851	1.0000
CorMatrix (theoretical)	1st	1.0000	-0.2194	-0.2541	1.0000	0.0000	0.0000
	2nd	-0.2194	1.0000	-0.8869	0.0000	1.0000	0.6554
	3rd	-0.2541	-0.8869	1.0000	0.0000	0.6554	1.0000

This table reports the summary statistics (sample mean, standard deviation, sample covariance matrix, and sample correlation coefficient matrix) of the filtered state variables for both SOCR and JCR. The table also presents the theoretical covariance and correlation coefficient matrices based on the parameter estimates given in Table 2.

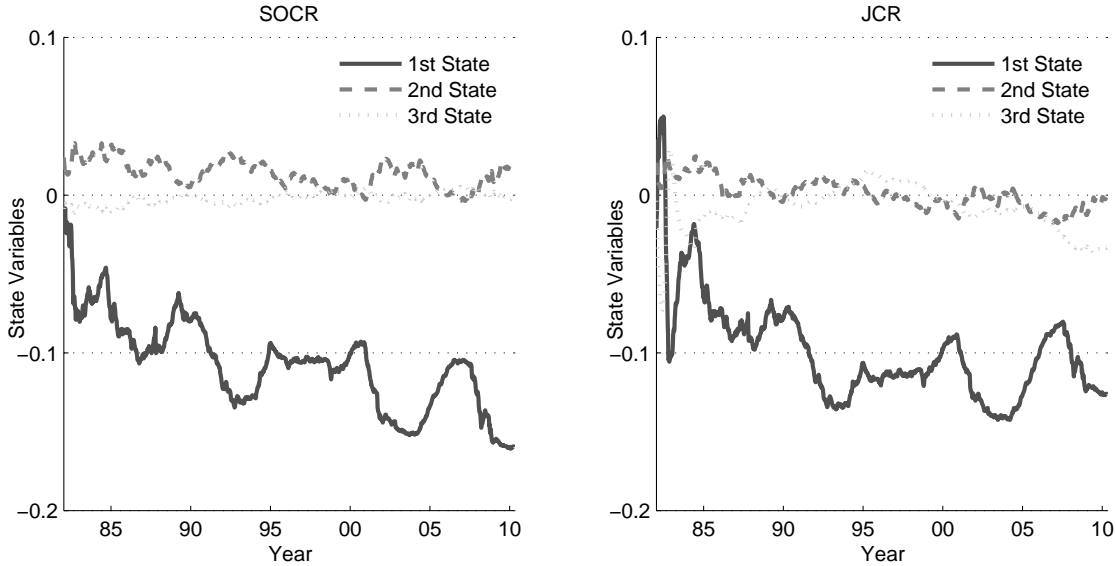
states. Although the states' dynamics look somewhat similar in SOCR and JCR as shown in Figure 4, their impacts on term structure of interest rates are quite different in the maturities range from 3-month to 10-year as shown in Figure 5. In SOCR, the first state variable almost has a unity impact on the term structure of interest rates from 3-month to 10-year, its factor loading only declines 7.4% over 10 years (from 1 at 3-month to 0.926 at 10-year). Therefore the first state variable clearly plays a role of "medium-run factor" which uniformly shifts the yield curve up to 10-year maturity.

The factor loading of the second state variable increases from 0.125 at 3-month to 3.445 at 10-year, exceeds that of the first state at 2-year, and declines after 18-year (not shown in Figure 5); the factor loading of the third state variable increases from 0.010 at 3-month to 4.888 at 10-year, exceeds that of the first state at 3-year, and declines after 20-year (not shown in Figure 5); they intersect at 2.21 around 5-year. And we will see in a minute that the first state variable well represents the short end of the term structure, and a linear combination of the second and third state variables can well explain the variation of the yield curve slope. In JCR, the factor loadings of the first and third state variables decline from 1 at 3-month to 0.595 and 0.429 at 10-year, respectively. The factor loading of the second state variable increases from 0.121 at 3-month, exceeds those of the first and third state variables at about 0.8 around 3-year, and reaches its peak at 1.473 around 7-year, then declines to 1.464 at 10-year. Therefore, in general, the dynamics of the short end of term structure are mainly controlled by the first and third state variables, the middle and long end of the term structure are more affected by the second state variables than by the other two.

In order to interpret the state variables economically, let's consider the relations between state variables and three important characteristics of the term structure of interest rates: a) curvature, b) slope, and c) short end. Following De Jong (2000), "curvature" is defined as two times the 1-year zero yield, minus the sum of the 3-month and 10-year zero yields; "slope" is the 10-year zero yield minus the 3-month zero yield; and "short end" is the 3-month zero yield. Table 5 reports the factor loadings of the three state variables for "curvature", "slope", and "short end" in both SOCR and JCR. In SOCR, we could see that the three states have very similar factor loadings with opposite signs for the curvature and slope, this is consistent with the empirical fact that the curvature and slope of yield curve are significantly negatively correlated. And for the short end, the factor loading of the first state well dominates

Figure 4: Dynamics of State Variables

This figure depicts the dynamics of state variables for both SOCR and JCR.



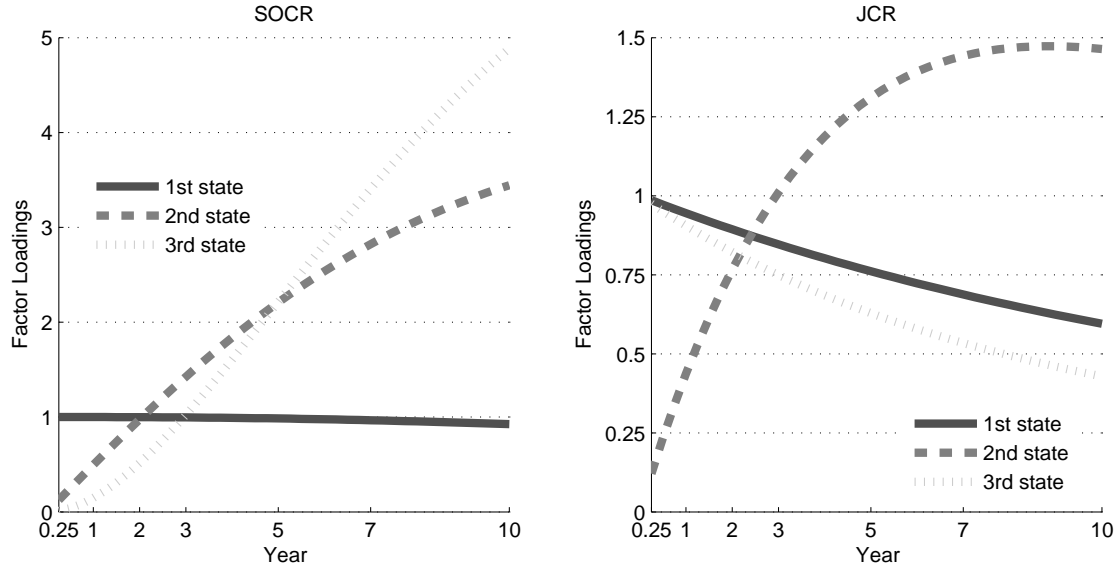
the other two. In JCR, we notice that the short end is almost completely determined by the first and third states. Table 5 also reports the VRs for the curvature, slope, and short end, they are 80.15%, 97.26%, and 99.46%, respectively. This shows that the model considered here is able to not only well capture the slope and level (slope plus short end) dynamics, but also explain a majority of the curvature variation. In the literature, the curvature variation is only expected to be captured by three-factor models (see, e.g., Litterman and Scheinkman, 1991; Dai and Singleton, 2000; De Jong, 2000). And the results documented here show that under our model specification a two-factor model can also capture the curvature dynamics quite well. The dynamics of the empirical and model implied curvature are presented in Figure 7(c).

Now we define two new variables: a) slp^{SOOCR} (short for the SOOCR “slope” factor) as

$$3.3197Z_t^{SOOCR} (2) + 4.8776Z_t^{SOOCR} (3);$$

Figure 5: Factor Loadings of State Variables

This figure depicts the term structures of factor loadings of state variables for both SOCR and JCR.



and b) shd^{JCR} (short for the JCR “short end” factor) as

$$0.9858Z_t^{JCR}(1) + 0.9751Z_t^{JCR}(3),$$

where $Z_t^{SOCR}(i)$ and $Z_t^{JCR}(i)$ are i th state variables of SOCR and JCR, respectively. Figure 6 presents the state variables’ (as well as their linear combinations’) contributions to the variations of the short end and slope. We can see from Figure 6(a), the first state variable in SOCR almost completely accounts for the short end variation, as its variance contribution is as high as 99%; from Figure 6(b), slp^{SOCR} , which is the linear combination of the second and third states of SOCR, explains about 85% slope variation; from Figure 6(c), shd^{JCR} , which is the linear combination of the first and third states of JCR, accounts for a great part (96%) of the short end variation. Panels (a) and (b) in Figure 7 confirm these observations.

Table 5: Factor Loadings for “Curvature”, “Slope”, and “Short End”

	SOCR			JCR			VR(%)
	1st	2nd	3rd	1st	2nd	3rd	
Curvature	0.0739	-2.5762	-4.6049	0.3087	-0.7103	0.4068	80.15
Slope	-0.0742	3.3197	4.8776	-0.3906	1.3432	-0.5466	97.26
Short End	1.0000	0.1249	0.0101	0.9858	0.1209	0.9751	99.46

This table reports the factor loadings of the three state variables for the “curvature”, “slope”, and “short end” in both SOCR and JCR. And the variance ratios (VRs) are also presented.

7.5 Short End as a “Medium-run Factor”

As shown above, the short end of the yield curve (3-month interest rate) can be exactly represented by the first state variable of SOCR. And if we look back at the factor loadings of this state in the left panel of Figure 5, we find a very flat factor loading term structure. Therefore the short end of the yield curve clearly plays a role of “medium-run factor” which has uniform impacts on dynamics of the yield curve up to 10-year maturity. This finding has an important implication for the short- and medium-term bond portfolios management, in the sense that an important systematic risk now becomes tradable. Therefore, by utilizing our model, relevant risk exposures of bond portfolios might be adjusted more precisely and quantitatively.

And this finding also helps us better understand the interactions between Macroeconomic policies and term structure of interest rates dynamics. As we all know, the U.S. Federal Reserve use open market operations for implementing monetary policy (Edwards, 1997). Specifically, these purchases and sales of U.S. Treasury and federal agency securities largely determine the federal funds rate—the interest rate at which depository institutions lend balances at the Federal Reserve to other depository in-

stitutions overnight. The federal funds rate, in turn, affects monetary and financial conditions, therefore, directly influences the short-term interest rates. In other words, the short end of the yield curve is highly sensitive to the macro-economic environment and monetary policy. So if the short end of the yield curve plays as a “medium-run factor” which uniformly shifts the yield curve, then the U.S. Federal Reserve might safely assume her monetary policies would have effective impacts on the U.S. economy as long as open market operations work.

8 Conclusion

Given the time-invariant hump shape volatility function as (1), the corresponding HJM model is guaranteed to admit *FDRs*. In this paper, working with the HJM model with volatility function of (1), a type of HJM based affine term structure model is theoretically derived. Specifically, Linear Realization Theory is applied to efficiently construct the *FDRs* via two canonical realization algorithms: Standard Observable Canonical Realization and Jordan Canonical Realization. And the commonly adopted method of constructing an *FDR*, which is set forth in Bhar and Chiarella (1996) for the first time, is generalized to an extent. It is shown, via a concrete example, that the resulting *FDR* is actually the Jordan Canonical Realization.

Three conclusions can be drawn from empirical analysis of the refined two-factor model. First, the model introduced in this paper provides great fit to the term structure data. In this aspect, the two-factor model of this type is much better than the traditional two-factor affine model, and comparable to the traditional three-factor affine model. However, the pricing errors exhibit relatively high autocorrelations. Exploring the economic information contained therein is an interesting topic for future research.

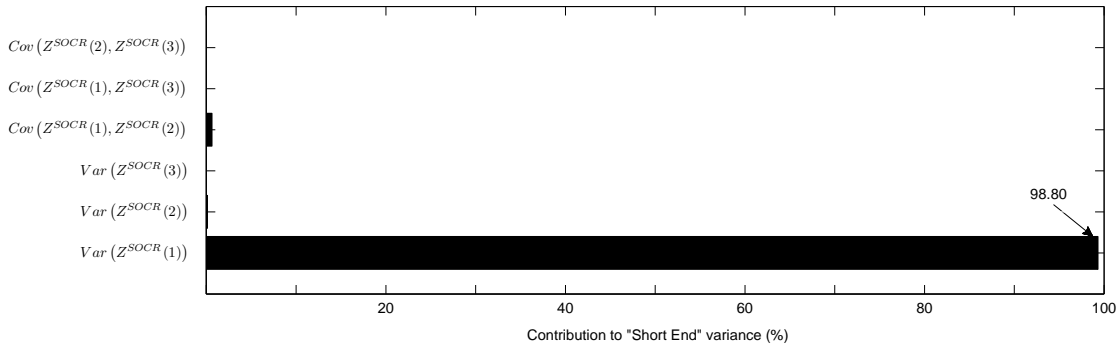
Second, the three state variables are closely linked to the short end and slope of the yield curve, and are able to well capture the curvature variation. Particularly, in SOCR, the first state variable almost completely represents the short end of the yield curve, and a linear combination of the second and third state variables well accounts for the slope of the yield curve; in JCR, a linear combination of the first and third state variables is also precisely related to the short end. Under our model specification, the curvature of the yield curve, which is only expected to be captured by three-factor models, could also be well explained using a two-factor model.

Last but not least, using SOCR, it is found that the short end of the yield curve can be regarded as a “medium-run factor” which uniformly shifts the yield curve. For bond portfolios management practitioners, this finding means that a systematic risk in bond market now becomes tradable. For macro-economic researchers, this finding suggests a new quantitative perspective from which the interactions between Macro-economic policies and term structure of interest rates dynamics can be understood more clearly and deeply.

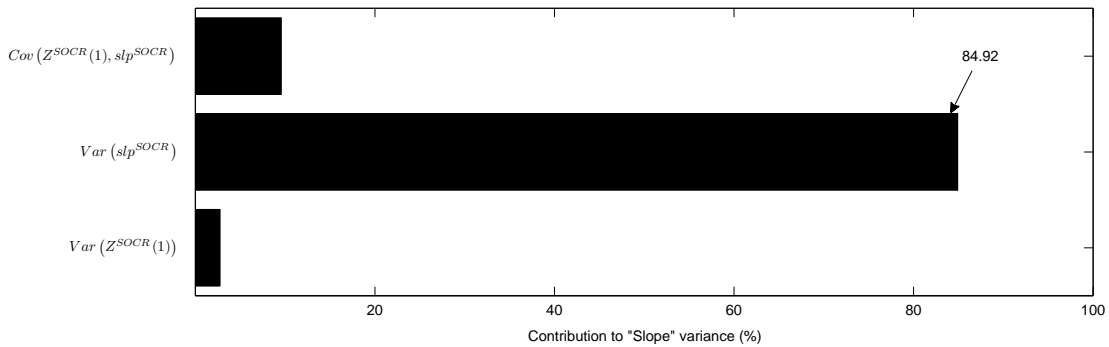
Figure 6: State Variables' Contributions to "Short End" and "Slope"

This figure depicts the state variables' (as well as their linear combinations') contributions to the variations of the short end and slope. Panel (a) shows the contribution of the first state variable of SOCR to the short end variation; Panel (b) shows slp^{SOCR} 's contribution to the slope variation in SOCR; Panel (c) shows shd^{JCR} 's contribution to the short end variation in JCR. Here "Var" and "Cov" denote the variance and covariance operators, respectively.

(a) Contributions to "Short End" in SOCR



(b) Contributions to "Slope" in SOCR



(c) Contributions to "Short End" in JCR

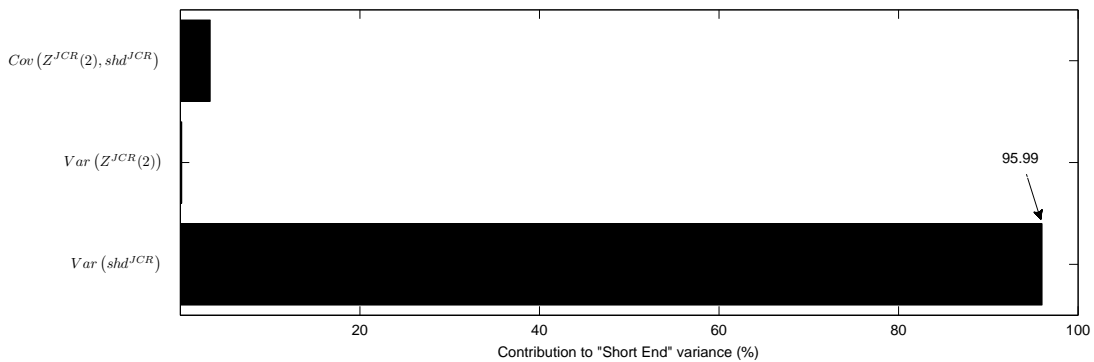
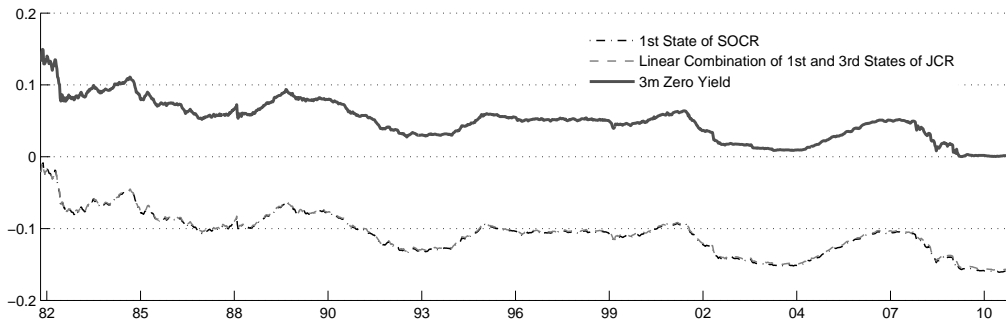


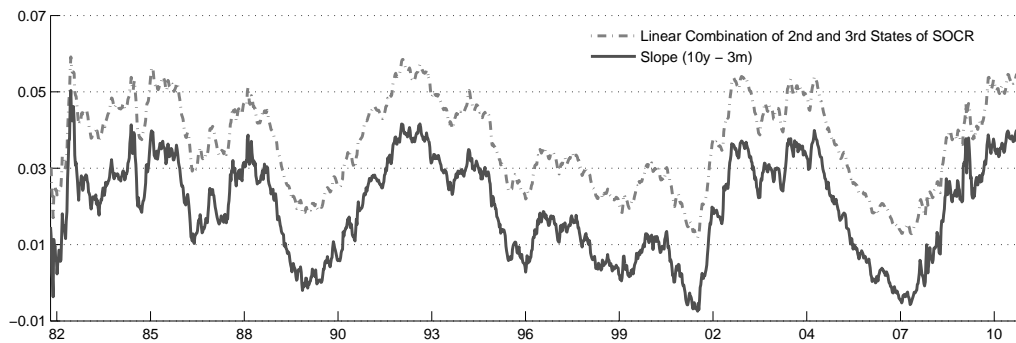
Figure 7: Dynamics of "Short End", "Slope", and "Curvature"

This figure depicts the dynamics of the short end, slope, and curvature of the term structure. For comparison, the dynamics of the first state of SOCR, slp^{SOCR} , shd^{JCR} , and model implied curvature are also presented in the panels. Panel (a) shows dynamics of the short end, 1st state of SOCR, and shd^{JCR} ; Panel (b) shows dynamics of the slope and slp^{SOCR} ; Panel (c) shows the dynamics of both empirical and model implied curvature.

(a) "Short End" v.s. 1st State of SOCR and shd^{JCR}



(b) "Slope" v.s. slp^{SOCR}



(c) "Empirical Curvature" v.s. "Model Implied Curvature"



Appendices

A The Refined Model

In the refined model, following volatility function is assumed,

$$\sigma(x) = [a_{10} \exp(-k_1 x), (a_{20} + a_{21} x) \exp(-k_2 x)]_{1 \times 2}.$$

The transfer function $G(s, 0) = \mathcal{L}[a_{10} \exp(-k_1 x), (a_{20} + a_{21} x) \exp(-k_2 x)](s)$ is given by

$$\begin{aligned} & \left[\frac{a_{10}}{k_1 + s'} \frac{a_{20}}{k_2 + s} + \frac{a_{21}}{(k_2 + s)^2} \right]_{1 \times 2} \\ &= \left[\frac{1}{k_1 + s'} \frac{1}{k_2 + s'} \frac{1}{(k_2 + s)^2} \right] \begin{bmatrix} a_{10} & 0 \\ 0 & a_{20} \\ 0 & a_{21} \end{bmatrix}. \end{aligned}$$

The realization algorithms are applied to

$$\left[\frac{1}{k_1 + s'} \frac{1}{k_2 + s'} \frac{1}{(k_2 + s)^2} \right]. \quad (\text{A.1})$$

- Standard Observable Canonical Realization

1. Expand (A.1) about $|s| = \infty$

$$\begin{aligned} & \left[\frac{1}{k_1 + s'} \frac{1}{k_2 + s'} \frac{1}{(k_2 + s)^2} \right] \\ &= [1, 1, 0] \frac{1}{s} + [-k_1, -k_2, 1] \frac{1}{s^2} + [k_1^2, -2k_1, k_2^2, -2k_2] \frac{1}{s^3} + \dots \end{aligned}$$

2. The Monic least common multiple of the denominators is given by

$$\begin{aligned}
 p(s) &= (k_1 + s)(k_2 + s)^2 \\
 &= k_1 k_2^2 + (k_2^2 + 2k_1 k_2)s + (2k_2 + k_1)s^2 + s^3 \\
 &\equiv p_0 + p_1 s + p_2 s^2 + s^3,
 \end{aligned}$$

so $n = 3$;

3. \mathbf{A} , \mathbf{B} , and $\mathbf{C}(x)$ can be computed as

$$\begin{aligned}
 \mathbf{A} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -k_1 k_2^2 & -(k_2^2 + 2k_1 k_2) & -(k_1 + 2k_2) \end{bmatrix}_{3 \times 3} \\
 \mathbf{B} &= \begin{bmatrix} 1 & 1 & 0 \\ -k_1 & -k_2 & 1 \\ k_1^2 & k_2^2 & -2k_2 \end{bmatrix}_{3 \times 3} \begin{bmatrix} a_{10} & 0 \\ 0 & a_{20} \\ 0 & a_{21} \end{bmatrix}_{3 \times 2} \\
 \mathbf{C}(x) &= \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \\
 &\quad \times \exp \left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -k_1 k_2^2 & -(k_2^2 + 2k_1 k_2) & -(k_1 + 2k_2) \end{bmatrix} x \right).
 \end{aligned}$$

- Jordan Canonical Realization

$$\begin{aligned}\tilde{\mathbf{A}} &= \begin{bmatrix} -k_1 & 0 & 0 \\ 0 & -k_2 & 0 \\ 0 & 1 & -k_2 \end{bmatrix} \\ \tilde{\mathbf{B}} &= \begin{bmatrix} a_{10} & 0 \\ 0 & a_{21} \\ 0 & a_{20} \end{bmatrix} \\ \tilde{\mathbf{C}}(x) &= \begin{bmatrix} e^{-xk_1} & xe^{-xk_2} & e^{-xk_2} \end{bmatrix}.\end{aligned}$$

B Kalman Filter

We start the KF by choosing the initial values of the state variables and their covariance matrix as their steady state values⁶

$$Z_{0|0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, P_{Z,0|0} = \mathbf{Cov}(Z_t).$$

Given $Z_{t-1|t-1}$ and $P_{Z,t-1|t-1}$, the *ex ante* prediction of the states and its covariance matrix are given by

$$\begin{aligned}Z_{t|t-1} &= \exp(\mathbf{A}h) Z_{t-1|t-1}, \\ P_{Z,t|t-1} &= \exp(\mathbf{A}h) P_{t-1|t-1} \exp(\mathbf{A}h)^\top + \mathbf{Cov}_h(Z_t),\end{aligned}$$

where

$$\mathbf{Cov}_h(Z_t) = \exp(\mathbf{A}h) \left[\int_0^h \exp(-\mathbf{A}s) \mathbf{B}\mathbf{B}^\top \exp(-\mathbf{A}s)^\top \mathbf{d}s \right] \exp(\mathbf{A}h)^\top.$$

⁶With a slight abuse of notation, here \mathbf{A} denotes both \mathbf{A} and $\tilde{\mathbf{A}}$, \mathbf{B} denotes both \mathbf{B} and $\tilde{\mathbf{B}}$, and $\mathbf{C}(\cdot)$ denotes both $\mathbf{C}(\cdot)$ and $\tilde{\mathbf{C}}(\cdot)$.

Given $Z_{t|t-1}$ and $P_{Z,t|t-1}$, the *ex ante* predictions of the measurement and the associated covariance become

$$y_{t|t-1} = \varphi + \frac{\int_0^x \Theta^*(s) \mathbf{d}s}{x} + \frac{\int_0^x \mathbf{C}(s) \mathbf{d}s}{x} Z_{t|t-1}.$$

$$P_{y_t} = \frac{\int_0^x \mathbf{C}(s) \mathbf{d}s}{x} P_{Z,t|t-1} \left[\frac{\int_0^x \mathbf{C}(s) \mathbf{d}s}{x} \right]^\top + \epsilon \mathbf{E} \mathbf{y} \mathbf{e}_8,$$

Finally, the *ex post* updates on the filtered states are given by

$$Z_{t|t} = Z_{t|t-1} + \Xi_t \left(y_t^x - y_{t|t-1} \right), \quad P_{Z,t|t} = P_{Z,t|t-1} - \Xi_t P_{y_t} \Xi_t^\top,$$

where $\Xi_t = P_{Z,t|t-1} \frac{\int_0^x \mathbf{C}(s) \mathbf{d}s}{x} P_{y_t}^{-1}$ is the Kalman gain.

Bibliography

- Bhar, R., Chiarella, C., 1996. Transformation of Heath-Jarrow-Morton Models to Markovian Systems. *The European Journal of Finance* 3, 1–26.
- Björk, T., Gombani, A., 1999. Minimal realizations of interest rate models. *Finance and Stochastics* 3, 413–432.
- Björk, T., Landén, C., 2002. On the construction of finite dimensional realizations for nonlinear forward rate models. *Finance and Stochastics* 6, 303–331.
- Björk, T., Svensson, L., 2001. On the Existence of Finite-Dimensional Realizations for Nonlinear Forward Rate Models. *Mathematical Finance* 11, 205–243.
- Bliss, R., Ritchken, P., Ronn, E., 1996. Empirical Tests of Two State-Variable Heath-Jarrow-Morton Models. *Journal of Money, Credit & Banking* 28, 452–476.
- Brace, A., Musiela, M., 1994. A MULTIFACTOR GAUSS MARKOV IMPLEMENTATION OF HEATH, JARROW, AND MORTON. *Mathematical Finance* 4, 259–283.
- Brockett, R., 1970. *Finite dimensional linear systems*. Wiley.
- Carverhill, A., 1994. When is the Short Rate Markovian? *Mathematical Finance* 4, 305–312.
- Chen, C.T., 1999. *Linear System Theory and Design*. 3rd. New York, NY: Oxford University Press.
- Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978.
- De Jong, F., 2000. Time series and cross-section information in affine term-structure models. *Journal of Business & Economic Statistics* 18, 300–314.

- De Jong, F., Santa-Clara, P., 1999. The Dynamics of the Forward Interest Rate Curve: A Formulation with State Variables. *Journal of Financial and Quantitative Analysis* 34, 131–157.
- Edwards, C., 1997. Open market operations in the 1990s. *Fed. Res. Bull.* 83, 859.
- Heath, D., Jarrow, R., Morton, A., 1992. Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica* 60, 77–105.
- Jarrow, R., Li, H., Zhao, F., 2007. Interest Rate Caps "Smile" Too! But Can the LIBOR Market Models Capture the Smile? *The Journal of Finance* 62, 345–382.
- Lemke, W., 2005. *Term Structure Modeling and Estimation in a State Space Framework*. Springer.
- Litterman, R., Scheinkman, J., 1991. Common factors affecting bond returns. *The Journal of Fixed Income* 1, 54–61.
- Mercurio, F., Moraleda, J., 2000. An analytically tractable interest rate model with humped volatility. *European Journal of Operational Research* 120, 205–214.
- Mercurio, F., Moraleda, J., 2001. A family of humped volatility models. *The European Journal of Finance* 7, 93–116.
- Ritchken, P., Chuang, I., 2000. Interest Rate Option Pricing With Volatility Humps. *Review of Derivatives Research* 3, 237–262.
- Ritchken, P., Sankarasubramanian, L., 1995. Volatility Structures of Forward Rates and the Dynamics of the Term Structure. *Mathematical Finance* 5, 55–72.
- Schutter, B., 2000. Minimal state-space realization in linear system theory: an overview. *Journal of Computational and Applied Mathematics* 121, 331–354.

Sercu, P., Xueping, W., 1997. The information content in bond model residuals: An empirical study on the Belgian bond market. *Journal of Banking and Finance* 21, 685–720.

Trolle, A., Schwartz, E., 2009. A general stochastic volatility model for the pricing of interest rate derivatives. *Review of Financial Studies* 22, 2007.

Turan, B., Heidari, M., Wu, L., 2009. Predictability of Interest Rates and Interest-Rate Portfolios. *Journal of Business & Economic Statistics* 27.