Bond Valuation Under a Discrete-Time Regime-Switching Term-Structure Model and its Continuous-Time Extension

Robert J. Elliott * Tak Kuen Siu† Alex Badescu‡

July 2, 2011

Abstract

We consider a discrete-time, Markov, regime-switching, affine term structure model for valuing bonds and other interest-rate securities. The proposed model incorporates the impact of structural changes in (macro)-economic conditions on interest rate dynamics. The market in the proposed model is, in general, incomplete. A modified version of the Esscher transform, namely, a double Esscher transform, is used to specify a price kernel so that both market and economic risks are taken into account. We derive a simple way to give exponential-affine forms of bond prices using backward induction. We also consider a continuous-time extension of the model and derive exponential-affine forms of bond prices using the concept of stochastic flows.

Keywords: Stochastic Interest Rates; Double Esscher Transform; Regime Switching Risk; Markov Chain; Exponential Affine Form; Continuous-Time Models; Product Density Processes; Stochastic flows.

*Haskayne School of Business, University of Calgary, Calgary, Alberta, Canada; School of Mathematical Sciences, University of Adelaide, Adelaide, South Australia, Australia

†Department of Applied Finance and Actuarial Studies, Faculty of Business and Economics, Macquarie University, Sydney, New South Wales, Australia

‡Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada

1 Introduction

Modeling the term structure of interest rates has long been an important topic in economics and finance. Numerous models for term structures of interest rates have been proposed in the literature. Early models for describing the stochastic behavior of interest rates are single-factor based. These include the models introduced by Merton (1973), Vasciek (1977), Cox et al. (1985), and others. The single-factor models have the advantage that closed-form pricing formulae for zero-coupon bonds are available. This is appealing from the practical perspective. However, the single-factor models imply that the long-term rate is a deterministic function of the spot rate, and that the prices of bonds with different maturities are perfectly correlated. These properties are not consistent with the empirical behavior of yield curves. Further, numerous empirical studies, (see, for example, Chapman and Pearson (2001)), reveal that the volatility structure of the short rate is more complex than those arising from the single-factor models. This motivates the quest for more flexible and general term structure models.

Multi-factor models provide a richer structure and more degrees of freedom with which to model different behavior of term structures of interest rates. They postulate that the evolution of the term structure of interest rates over time depends on several, (macro-economic), factors, and hence, yields predicted from these models are functions of these factors. Some multi-factor models include those of Brennan and Schwartz (1979), Longstaff and Schwartz (1992), and others. Brennan and Schwartz (1979) developed a multi-factor model based on a two-dimensional diffusion model for two state variables representing the short rate and the long rate, where the long rate is represented as the reciprocal of a consol's price. However, Dybvig et al. (1996) showed theoretically that the long rate is non-decreasing, so a diffusion model is not suitable for modeling the long rate. Chen (1996) proposed a three-factor model with factors including the short rate, the short-term mean rate and the short rate volatility. Duffie and Kan (1996) proposed a general multi-factor affine term structure model, which nests many of the single-factor and multi-factor models in the existing literature.

Many existing stochastic term structure models were primarily developed for the purposes of modeling short or medium-term behavior of the term structures of interest rates and the pricing, hedging and risk management of some short or medium-term interest-rate products. However, for modeling the long-term behavior of interest rates and managing the risk of long-term interest-rate related products including insurance products, it is important to model the long-term behavior of interest rates. Over a longer time period, there may be structural changes in the term structure of interest rates, which can be attributed to the structural changes of macro-economic conditions, economic fundamentals and monetary policies. It is important to capture these structural changes in modeling the long-term behavior of interest rates. Regime-switching models provide a natural way to model such structural changes. Indeed, some empirical studies indicate that switching behavior is present in interest rates data and that the regime-switching models provide a good fit for this data, (see, for example, Roma and Torous (1997) and Ang and Bekaert (2002a, b)).

An important application of term structure models is the valuation of interest-rate instruments, such as bonds and fixed income derivatives. Ideally, it is desirable to have a term structure model, which is flexible and general enough to describe various empirical features of interest rate data and is tractable for these valuations. Duffie and Kan (1996) studied the relationship between a parametric form of interest rate models and the exponential-affine form of bond price. Elliott and van der Hoek (2001) established the relationship between the parametric form of interest rate models and the exponential-affine form of the bond price by exploiting the use of stochastic flows and forward measures. They constructed explicitly the bond price using linear ordinary differential equations. Elliott and Siu (2009) extended some results of Elliott and van der Hoek (2001) and derived a Markov-modulated, exponential-affine, bond price formula under both a Markov, regime switching, Hull-White model and a Markov, regime switching, Cox-Ingersoll-Ross (CIR) model. Other works on continuous-time, regime switching, short rate models include Naik and Lee (1997), Dai and Singleton (2003), Elliott and Mamon (2003a, b), Elliott and Kopp (2004), and Elliott and Wilson (2007). These models can be classified as affine term structure models. Much of the literature about bond valuation in affine term structure models focuses on continuous-time models. However, there are fewer contributions to bond valuation for discrete-time affine term structure models, especially when a regime switching effect is present. For econometric purposes, discrete-time term structure models seem more easy to implement and estimate than their continuous-time counterparts. Some works include Evans (1998), Dai and Singleton (2000), and Ang and Bekaert (2002c), and others. The typical approach seems to model the short rate dynamics as a single-factor regime switching, model. Many empirical studies reveal that the single-factor, regime-switching short rate model outperforms its constant-coefficient counterparts when fitting real interest rate data. Since richer specifications can provide more flexible structures for incorporating more stylized features exhibited by yield curve dynamics, some extensions to the single-factor, regime-switching model have been proposed in the literature. For example, Bansal and Zhou (2002) proposed a term structure model where both the short rate dynamics and the market price of risk are modeled by discrete-time Gaussian regime switching processes. More recently, Dai *et al.* (2007) empirically examined a discrete time, regime switching model where the short rate follows a three-factor Gaussian model with state-dependent market prices of risk. Ang *et al.* (2007) employed a three-factor representation for yield curves by incorporating regime changes for the inflation factor and the term structure factor, and assuming a regime-invariant price of risk for the factors. Although multi-factor models may provide more flexibility in incorporating stylized features of yield curve dynamics, we focus on the case when only the short rate dynamics are modeled. However, the pricing method proposed here can also be extended to richer model dynamics.

In this paper we introduce a discrete-time, Markov, regime-switching, affine term structure model for valuing bonds and other interest-rate securities. Under this model, the set of model parameters in force at a particular time depends on the state of the economy at that time. This set of model parameters switches to another set when there is a transition in the state of the economy. Here the stochastic evolution of the state of the economy is modeled by a discrete-time, finite-state, Markov chain. Consequently, the model incorporates the impact of structural changes in (macro)-economic conditions on interest rate dynamics. It also has some econometric advantages. In particular, the proposed, discrete-time, model is easier to estimate than its continuous-time counterparts. The market containing the proposed model is, in general, incomplete so there is more than one price kernel, or stochastic discount factor, for valuation. A key question is how to determine a price kernel. Here we employ a modified version of a time-honored tool in actuarial science, namely, the Esscher transform, to specify a price kernel. The modified version is called a double Esscher transform and is defined by the product of two density processes for measure changes, one for a measure change for the interest rate process and one for a measure change for the Markov chain describing the state of the economy. The rationale of introducing the double Esscher transform to specify a price kernel is to take into account both the market risk due to fluctuations of interest rate and the economic risk due to transitions of the state of the economy. In addition to providing econometric advantages, the proposed, discrete-time, model also simplifies the bond valuation procedure compared with continuous-time models. We derive exponential-affine forms of bond prices by backward induction. The derivation procedure is more simple than its continuous-time counterpart where techniques from continuous-time stochastic calculus, such as stochastic flows, are used. We also consider a continuous-time extension of the model. The continuous-time model provides further theoretical justifications for the proposed discrete-time model. A pricing kernel is specified by the product of two density processes, one for a measure change for the standard Brownian motion and another for a measure change for a continuous-time, finite-state, Markov chain. A Girsanov transform for a Markov chain is used for the measure change. The concept of stochastic flows is then used to derive exponential-affine forms of bond prices.

This paper is structured as follows. The next section presents the discrete-time, Markov, regime-switching short rate model. In Section 3, we describe the double Esscher transform and its use for specifying a price kernel for valuation. We derive the exponential-affine forms of bond prices in Section 4. The continuous-time extension of the proposed model is considered in Section 5. We also discuss the specification of the pricing kernel based on the product of two density processes. In Section 6, we derive an exponential formula for the bond price using the concept of stochastic flows. The final section summarizes the paper.

2 The Regime Switching Short Rate Model

Consider a discrete-time economy with finite time horizon and time index set $\mathcal{T} = \{k | k = 0, 1, 2, \cdots, T\}$, where T is a positive integer and $T < \infty$. To model uncertainty, we adopt a complete probability space $(\Omega, \mathcal{F}, P)$, where P is a real-world probability. We assume that this probability space is rich enough to incorporate uncertainties due to fluctuations of interest rate and transitions of the states of an economy.

We assume there are N distinct states of the economy and the stochastic evolution of the state of the economy is governed by a discrete-time, N-state Markov chain $\mathbf{X} := \{\mathbf{X}_k\}_{k \in \mathcal{T}}$. Without the loss of generality, we take the state space of the chain $\mathbf{X}$ as the set of canonical unit column vectors $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N\}$, where $\mathbf{e}_i$ has unity in the i^{th} position and zero elsewhere. This is a mathematically convenient way to represent the state space of a Markov chain and is given in Elliott et al. (1994). To describe the statistical properties of the Markov chain, we need to introduce a time-dependent transition probability matrix, denoted as $\mathbf{A}_k := \{a_{jik}\}_{i,j=1,2,\cdots,N}$, under the measure P, where

$$a_{jik} = P(\mathbf{X}_{k+1} = \mathbf{e}_j | \mathbf{X}_k = \mathbf{e}_i) \ .$$

Here a_{jik}, $i, j = 1, 2, \cdots, N$, are state-dependent under the physical measure P. They also satisfy $\sum_{j=1}^{N} a_{jik} = 1$ and $a_{jik} \geq 0$.

Let $\mathcal{F}^{\mathbf{X}} := \{\mathcal{F}^{\mathbf{X}}_k\}_{k \in \mathcal{T}}$ denote the P-augmented filtration generated by the history of the chain $\mathbf{X}$, where $\mathcal{F}^{\mathbf{X}}_k$ is the P-augmented σ-field generated by the history of the Markov chain up to and including time k. $\mathcal{F}^{\mathbf{X}}_k$ contains information about possible histories of the states of the economy up to and including time k. Then Elliott *et al.* (1994), (see Chapter 2, Page 17, Equation (2.5) therein), derived the semimartingale dynamics for the Markov chain as:

$$\mathbf{X}_{k+1} = \mathbf{A}_k \mathbf{X}_k + \mathbf{M}_{k+1} \ , \quad k = 0, 1, \cdots, T-1 \ . \tag{1}$$

Here $\mathbf{M} := \{\mathbf{M}_k\}_{k=1,2,\cdots,T}$ is an $\Re^N$-valued martingale increment process so that

$$E[\mathbf{M}_{k+1} | \mathcal{F}^{\mathbf{X}}_k] = 0 \ .$$

Let $\mathbf{Y} := \{\mathbf{Y}_k\}_{k \in \mathcal{T}} = (Y_{1k}, Y_{2k}, \cdots, Y_{Lk})'$ denote the L-factor state process. Write $\mathcal{F}^{\mathbf{Y}} := \{\mathcal{F}^{\mathbf{Y}}_k\}_{k \in \mathcal{T}}$ for the P-augmented filtration generated by the factor process $\mathbf{Y}$; that is, for each $k \in \mathcal{T}$, $\mathcal{F}_k^{\mathbf{Y}}$ is the information set generated by the history of the state process up to and including time k. We suppose that economic agents observe the current and historical information about both the state process and the economic regimes. This information set is denoted by $\mathcal{G}_k = \mathcal{F}_k^{\mathbf{Y}} \vee \mathcal{F}_k^{\mathbf{X}}$, for any $k \in \mathcal{T}$.

Let $\langle \cdot, \cdot \rangle$ denote the scalar product in $\Re^N$. For each $k = 1, 2, \cdots, T$, we define the following regime-dependent vector $\boldsymbol{\alpha}(\mathbf{X}_k) := (\langle \boldsymbol{\alpha}_1, \mathbf{X}_k \rangle, \langle \boldsymbol{\alpha}_2, \mathbf{X}_k \rangle, \cdots, \langle \boldsymbol{\alpha}_L, \mathbf{X}_k \rangle)'$ and matrices $\boldsymbol{\Gamma}(\mathbf{X}_k) := (\langle \boldsymbol{\Gamma}_{ij}, \mathbf{X}_k \rangle)_{1 \le i,j \le L}$ and $\boldsymbol{\Sigma}(\mathbf{X}_k) := (\langle \boldsymbol{\Sigma}_{ij}, \mathbf{X}_k \rangle)_{1 \le i,j \le L}$, where $\boldsymbol{\alpha}_i, \boldsymbol{\Gamma}_{ij}$ and $\boldsymbol{\Sigma}_{ij}$ are $N \times 1$ real-valued vectors for any $1 \le i, j \le L$. Note that the scalar product $\langle \cdot, \cdot \rangle$ selects the component of these vectors that is in force at time k according to the state of the chain $\mathbf{X}_k$ at time k. So the dependence on $\mathbf{X}_k$ comes from the scalar product.

In order to fully specify the mean of the state process under the real-world probability measure P, let $\boldsymbol{\theta}_k(\mathbf{X}_k) = (\langle \boldsymbol{\theta}_{1k}, \mathbf{X}_k \rangle, \langle \boldsymbol{\theta}_{2k}, \mathbf{X}_k \rangle, \cdots, \langle \boldsymbol{\theta}_{Lk}, \mathbf{X}_k \rangle)'$ where $\boldsymbol{\theta}_{ik}$ is an $N \times 1$ vector of $\mathcal{G}_k$-measurable random variables for any $1 \le i \le L$ representing the regime dependent market price of factor risk. For example, when the state of the economy is in state $\mathbf{e}_j$ then the j^{th} component of $\boldsymbol{\theta}_{ik}$, denoted by θ_{ijk}, is the market price of factor i risk in state j at time k.

Suppose $\boldsymbol{\varepsilon} := \{\boldsymbol{\varepsilon}_k\}_{k=1,2,\cdots,T}$ denotes a sequence of $N \times 1$ independent and identically distributed random variable with $\boldsymbol{\varepsilon}_k \sim N(\mathbf{0}, \mathbf{I})$. We further assume that $\boldsymbol{\varepsilon}$ and $\mathbf{X}$ are stochastically independent, which is the usual standard assumption in most of the financial econometrics literature with regime switching modeling.

We assume that under the real-world probability measure P the evolution of the factor process is governed by the following discrete-time, Markov, regime-switching, first-order, autoregressive time series model:

$$\mathbf{Y}_{k+1} = \boldsymbol{\alpha}(\mathbf{X}_k) + \boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k)\boldsymbol{\theta}_k(\mathbf{X}_k) + \boldsymbol{\Gamma}(\mathbf{X}_k)\mathbf{Y}_k + \boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\varepsilon}_{k+1} \tag{2}$$

Let $r := \{r_k\}_{k \in \mathcal{T}}$ denote the process of short term interest rate. We assume the dynamic of r_k is regime dependent and is given by the following affine function of the factor process $\mathbf{Y}_k$:

$$r_k := r_k(\mathbf{X}_k) = d_0(\mathbf{X}_k) + \mathbf{d}_1'(\mathbf{X}_k)\mathbf{Y}_k \tag{3}$$

Here $d_0(\mathbf{X}_k) = \langle \mathbf{d}_0, \mathbf{X}_k \rangle$ and $\mathbf{d}_1(\mathbf{X}_k) = (\langle \mathbf{d}_{11}, \mathbf{X}_k \rangle, \langle \mathbf{d}_{12}, \mathbf{X}_k \rangle, \cdots, \langle \mathbf{d}_{1L}, \mathbf{X}_k \rangle)'$, where $\mathbf{d}_0$ and $\mathbf{d}_{1i}$ are $N \times 1$ real-valued vectors for any $1 \leq i \leq L$.

Similar discrete-time structures for the factor process and interest rate dynamics have been considered for estimation purposes by Dai *et al.* (2007) and Ang *et al.* (2007). However, our proposed model from equations (2) and (3) is more flexible since some of their assumptions are not imposed here. For example, in order to obtain a closed-form exponential affine form for bond prices, both studies require the coefficient of $\mathbf{Y}_k$ in the factor process dynamics to be constant. We relax this assumption by assuming that $\mathbf{\Gamma}$ is regime dependent. Another crucial assumption imposed by Dai *et al.* (2007) to facilitate bond pricing is to let the "loading" $\mathbf{d}_1$ to have the same value across regimes, while we assume $\mathbf{d}_1$ to depend on the Markov chain as well. A similar condition was considered in Bansal and Zhou (2002).

A common approach in the derivative pricing literature is to model the market price of risk under the real-world probability measure P. Following this line,we assume that the factor and regime dependent price of risk $\boldsymbol{\theta}_k$ is present in the dynamics of $\mathbf{Y}_k$ in (2). However, this price of risk will be eliminated via risk neutralization; that is, it does not appear in the risk-neutral dynamics of the short term interest rate to be derived later. A similar approach was adopted in Dai *et al.* (2007) for the model specification of the short term interest rate under a real-world probability measure. There are at least **two** advantages of considering the specification of short term interest rate dynamics given here. Firstly, under this specification, we can derive a tractable, exponential, affine form for the bond price. Secondly, this specification simplifies the numerical procedure for the practical implementation of the exponential affine form of the bond price. The estimation procedure for the model parameters under the real-world probability measure P can be further simplified if we assume that $\boldsymbol{\theta}_k$ depends only on the market regime $\mathbf{X}_k$ at time k, (i.e. $\boldsymbol{\theta}_k$ does not depend on time, or $\boldsymbol{\theta}_k = \boldsymbol{\theta}$ for each $k \in \mathcal{T} \backslash \{0\}$, where $\boldsymbol{\theta}$ is a constant vector in $\Re^N$).

In the proposed model, given the information set up to time k, $\mathcal{G}_k$, the state process under the real-world measure P is conditionally multivariate normal distributed with regime-dependent L-dimensional mean vector and $L \times L$ dimensional covariance matrix:

$$\begin{aligned}
\mathbf{m}_k(\mathbf{X}_k) &:= E^P[\mathbf{Y}_{k+1}|\mathcal{G}_k] = \boldsymbol{\alpha}(\mathbf{X}_k) + \boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k)\boldsymbol{\theta}_k(\mathbf{X}_k) + \boldsymbol{\Gamma}(\mathbf{X}_k)\mathbf{Y}_k \\
\mathbf{v}_k(\mathbf{X}_k) &:= Var^P[\mathbf{Y}_{k+1}|\mathcal{G}_k] = \boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k) \ .
\end{aligned}$$

3 Valuation of Zero-Coupon Bond by the Double Esscher Transform

The Esscher transform is a time-honored tool in actuarial science. Its history can be dated back to the seminal work of Esscher (1932), where the Esscher transform was first introduced to the actuarial science literature and was applied to approximate the distribution of the aggregate claims amount. It was then applied to various areas in actuarial science, such as premium calculation. Indeed, the seminal works of Bühlmann (1980, 1984), established the link between an economic premium principle in a pure risk exchange economy and the premium principle based on the Esscher transform when the (re)insurers have exponential utilities.

Gerber and Shiu (1994) pioneered the use of the Esscher transform for option valuation in a general asset price model. Indeed, the Esscher transform is a convenient tool for option valuation and its use for option valuation can be justified by some economic arguments based on the maximization of the expected power utility, (see Gerber and Shiu (1994)). The work of Gerber and Shiu (1994) highlighted the interplay between financial and actuarial pricing. This is an important research area in both actuarial research and financial mathematics research. Since the work of Gerber and Shiu, there has been much interest in the use of the Esscher transform for valuation. Bühmann et al. (1996) considered the use of the Esscher transform for asset pricing in a general semi-martingale asset price model. Bühmann et al. (1998) investigated the use of the Esscher transform for asset pricing in a discrete financial model and introduced the concept of the conditional Esscher transform which generalizes the Esscher transform for stochastic processes. Chan (1999) considered the use of the Esscher transform and other related methods for option valuation in a continuous-time geometric Lévy model. Yao (2001) explored the use of the Esscher transform to develop a general asset pricing model with a stochastic interest rate and a foreign exchange rate. Siu et al. (2004) introduced the use of the conditional Esscher transform for option valuation under a general GARCH model with non-normal innovations. Elliott et al. (2005) introduced the use of the regime-switching Esscher transform for option valuation under a regime-switching environment. The method was then considered in Siu (2005) for the valuation of liabilities underlying participating life insurance products.

In this section we present an extension of the Esscher transform, namely, a double Esscher transform, to determine a price kernel, or stochastic discount factor, for valuing a zero-coupon bond and other interest-rate related securities. The double Esscher transform is given by the product of two Esscher transforms, one defining the density process of a measure change for the short rate process and another defining the density process of a measure change for the Markov chain describing economic regimes. The idea of adopting the product of two density processes for measure changes of a diffusion process and a Markov chain was considered in Elliott and Siu (2010) in the context of a risk minimizing portfolio selection problem in a regime-switching environment. The rationale of using the product of two density processes is to provide a flexible framework which can incorporate both the financial risk and the economic risk in the portfolio risk minimization problem.

For valuing the bond and other interest-rate related securities, we need to specify a price kernel or a stochastic discount factor. Since the market in the proposed model here is incomplete, there is more than one price kernel or stochastic discount factor one can use for valuation. The key question is how one should select a price kernel for valuation. Of course, the choice of a price kernel is dictated by some factors, such as personal judgement, economic reasons or arguments, tractability and others. The use of the double Esscher transform here allows us to price both the risks due to fluctuations of market rates and transitions of economic regimes. It is of particular importance to price these two sources of risk when one prices long-term interest-rate products, such as bonds with long-maturities and insurance liabilities sensitive to interest rate movements. The double Esscher transform is also a convenient tool to value bonds and other interest-rate sensitive securities in a discrete-time, regime-switching, environment.

Following the methodology in Elliott and Siu (2010), we construct a new probability measure based on the product of two density processes, one for a measure change for the factor process and another for a measure change for the Markov chain. The two density processes are determined by two Esscher transforms, one for the exponential tilting of the interest rate process and another for the exponential tilting of the Markov chain.

Consider a G-adapted process $\lambda^{\boldsymbol{\theta}} := \{\lambda^{\boldsymbol{\theta}}_t\}_{t=1,2,\cdots,T}$ defined by setting:

$$\lambda^{\boldsymbol{\theta}}_t := \frac{\exp(-\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\mathbf{Y}_t)}{M_{\mathbf{Y}_t|\mathcal{G}_{t-1}}(-\boldsymbol{\theta}_{t-1}(\mathbf{X}_{t-1}))} \ .$$

Here $M_{\mathbf{Y}_t|\mathcal{G}_{t-1}}(\cdot)$ is the conditional moment generating function of $\mathbf{Y}_t$ given $\mathcal{G}_{t-1}$ under P; that is,

$$\begin{aligned}
& M_{\mathbf{Y}_t|\mathcal{G}_{t-1}}(-\boldsymbol{\theta}_{t-1}(\mathbf{X}_{t-1})) \\
= \ & E[\exp(-\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\mathbf{Y}_t)|\mathcal{G}_{t-1}] \\
= \ & E[\exp(-\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})(\mathbf{m}_{t-1}(\mathbf{X}_{t-1}) + \boldsymbol{\Sigma}(\mathbf{X}_{t-1})\boldsymbol{\varepsilon}_t))]|\mathcal{G}_{t-1}\} \\
= \ & \exp\left[-\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\Big(\boldsymbol{\alpha}(\mathbf{X}_{t-1}) + \boldsymbol{\Gamma}(\mathbf{X}_{t-1})\mathbf{Y}_{t-1} + \frac{1}{2}\boldsymbol{\Sigma}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}'(\mathbf{X}_{t-1})\boldsymbol{\theta}_{t-1}(\mathbf{X}_{t-1})\Big)\right] \ .
\end{aligned}$$

Thus, it is easy to see that

$$\lambda^{\boldsymbol{\theta}}_t = \exp\left(-\frac{1}{2}\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}'(\mathbf{X}_{t-1})\boldsymbol{\theta}_{t-1}(\mathbf{X}_{t-1}) - \boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}(\mathbf{X}_{t-1})\boldsymbol{\varepsilon}_t\right) \ .$$

Then the Esscher transform for the interest rate process is defined by a G-adapted process $\Lambda^{\boldsymbol{\theta}} = \{\Lambda^{\boldsymbol{\theta}}_k\}_{k\in\mathcal{T}}$ as follows:

$$\begin{aligned}
\Lambda^{\boldsymbol{\theta}}_0 &= 1 \ , \\
\Lambda^{\boldsymbol{\theta}}_k &:= \prod_{t=1}^{k}\lambda^{\theta}_t \\
&= \exp\left(-\frac{1}{2}\sum_{t=1}^{k}\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}'(\mathbf{X}_{t-1})\boldsymbol{\theta}_{t-1}(\mathbf{X}_{t-1}) - \sum_{t=1}^{k}\boldsymbol{\theta}'_{t-1}(\mathbf{X}_{t-1})\boldsymbol{\Sigma}(\mathbf{X}_{t-1})\boldsymbol{\varepsilon}_t\right) \ .
\end{aligned}$$

It is straightforward to show that $\Lambda^{\boldsymbol{\theta}}$ is a $(\mathcal{G}, P)$-martingale. The process $\Lambda^{\boldsymbol{\theta}}$ is the density process for the measure change of the interest rate process. Indeed, the process $\Lambda^{\boldsymbol{\theta}}$ is a discrete-time version of the exponential martingale which is used to specify the Radon-Nikodym derivative for Girsanov's change of measures in a continuous-time setting.

We shall now define the Esscher transform for the Markov chain. For each $t \in \mathcal{T}$, let $\boldsymbol{\Delta}_t := (\boldsymbol{\delta}_{ijt})_{1\leq i,j\leq N}$, be an $(N \times N)$-matrix of $\mathcal{G}_t$-measurable components. Write $\boldsymbol{\delta}_t = \boldsymbol{\Delta}_t\mathbf{X}_t \in \Re^N$.

Consider a G-adapted process $\lambda^{\boldsymbol{\delta}} := \{\lambda_t^{\boldsymbol{\delta}}\}_{t=1,2,\cdots,T}$ defined by setting:

$$\lambda_t^{\boldsymbol{\delta}} := \frac{\exp\big(\langle \boldsymbol{\delta}_{t-1}, \mathbf{X}_t\rangle\big)}{M_{\mathbf{X}_t|\mathcal{G}_{t-1}}(\boldsymbol{\delta}_{t-1})} \ . \tag{4}$$

Here $M_{\mathbf{X}_t|\mathcal{G}_{t-1}}(\boldsymbol{\delta}_{t-1})$ is the conditional moment generation function of $\mathbf{X}_t$ given $\mathcal{G}_{t-1}$ under the measure P evaluated at $\boldsymbol{\delta}_{t-1}$; that is,

$$M_{\mathbf{X}_t|\mathcal{G}_{t-1}}(\boldsymbol{\delta}_{t-1}) := E[\exp\big(\langle \boldsymbol{\delta}_{t-1}, \mathbf{X}_t\rangle\big)|\mathcal{G}_{t-1}] \ . \tag{5}$$

As $\boldsymbol{\delta}_{t-1} := \boldsymbol{\Delta}_{t-1}\mathbf{X}_{t-1}$, $t = 1, 2, \cdots, T$, we can define

$$\lambda_{ij,t-1} := \exp(\langle \boldsymbol{\Delta}_{t-1}\mathbf{e}_j, \mathbf{e}_i\rangle) = \exp(\delta_{ij,t-1}) \ , \quad i, j = 1, 2, \cdots, N \ ,$$

and write $\mathbf{L}_{t-1} := (\lambda_{ij,t-1})_{i,j=1,2,\cdots,N}$, where $\lambda_{ij,t-1} > 0$. Then the numerator of $\lambda_t^{\boldsymbol{\delta}}$ has a bilinear form in $\mathbf{X}_{t-1}$ and $\mathbf{X}_t$; that is,

$$\exp(\langle \boldsymbol{\delta}_{t-1}, \mathbf{X}_t\rangle) = \mathbf{X}'_{t-1}\mathbf{L}'_{t-1}\mathbf{X}_t \ .$$

So

$$\begin{aligned} M_{\mathbf{X}_t|\mathcal{G}_{t-1}}(\boldsymbol{\delta}_{t-1}) &= E[\mathbf{X}'_{t-1}\mathbf{L}'_{t-1}\mathbf{X}_t|\mathcal{G}_{t-1}] \\ &= \mathbf{X}'_{t-1}(\mathbf{L}'_{t-1}\mathbf{A}_{t-1})\mathbf{X}_{t-1} \ . \end{aligned} \tag{6}$$

This is a quadratic form in $\mathbf{X}_{t-1}$.

Consequently, from the semi-martingale dynamics (1) of the chain $\mathbf{X}$,

$$\begin{aligned} \lambda_t^{\boldsymbol{\delta}} &= \frac{\mathbf{X}'_{t-1}\mathbf{L}'_{t-1}\mathbf{X}_t}{\mathbf{X}'_{t-1}(\mathbf{L}'_{t-1}\mathbf{A}_{t-1})\mathbf{X}_{t-1}} \\ &= \frac{\mathbf{X}'_{t-1}\mathbf{L}'_{t-1}(\mathbf{A}_{t-1}\mathbf{X}_{t-1} + \mathbf{M}_t)}{\mathbf{X}'_{t-1}(\mathbf{L}'_{t-1}\mathbf{A}_{t-1})\mathbf{X}_{t-1}} \\ &= 1 + \frac{\mathbf{X}'_{t-1}\mathbf{L}'_{t-1}\mathbf{M}_t}{\mathbf{X}'_{t-1}(\mathbf{L}'_{t-1}\mathbf{A}_{t-1})\mathbf{X}_{t-1}} \ . \end{aligned} \tag{7}$$

We then define the Esscher transform for the Markov chain $\mathbf{X}$ by considering the following G-adapted process $\Lambda^{\boldsymbol{\delta}} := \{\Lambda_k^{\boldsymbol{\delta}}\}_{k\in\mathcal{T}}$:

$$\begin{aligned} \Lambda_0^{\boldsymbol{\delta}} &= 1 \ , \\ \Lambda_k^{\boldsymbol{\delta}} &:= \prod_{t=1}^{k} \lambda_t^{\boldsymbol{\delta}} = \prod_{t=1}^{k}\left(1 + \frac{\mathbf{X}'_{t-1}\mathbf{L}'_{t-1}\mathbf{M}_t}{\mathbf{X}'_{t-1}(\mathbf{L}'_{t-1}\mathbf{A}_{t-1})\mathbf{X}_{t-1}}\right) \ . \end{aligned} \tag{8}$$

By definition, we see that $\Lambda^{\boldsymbol{\delta}}$ is a $(\mathcal{G}, P)$-martingale.

Finally, we define the $(\mathcal{G}, P)$-adapted process $\Lambda^{\boldsymbol{\theta},\boldsymbol{\delta}} := \{\Lambda_k^{\boldsymbol{\theta},\boldsymbol{\delta}}\}_{k\in\mathcal{T}}$ by the product of the density processes $\Lambda^{\boldsymbol{\theta}}$ and $\Lambda^{\boldsymbol{\delta}}$ as:

$$\Lambda_k^{\boldsymbol{\theta},\boldsymbol{\delta}} = \Lambda_k^{\boldsymbol{\theta}} \cdot \Lambda_k^{\boldsymbol{\delta}} . \tag{9}$$

We then have the following lemma:

Lemma 3.1 $\Lambda^{\boldsymbol{\theta},\boldsymbol{\delta}}$ *is a* $(\mathcal{G}, P)$*-martingale.*

Proof Clearly $\Lambda_0^{\boldsymbol{\theta},\boldsymbol{\delta}} = 1$ by definition. For each $t \in \mathcal{T} \setminus \{0\}$, using the law of iterated expectations and stochastic independence between $\mathbf{X}$ and ε, we have:

$$\begin{aligned}
& E[\Lambda_k^{\boldsymbol{\theta},\boldsymbol{\delta}}|\mathcal{G}_{k-1}] \\
= \ & E[E(\Lambda_k^{\boldsymbol{\theta}} \cdot \Lambda_k^{\boldsymbol{\delta}}|\mathcal{G}_{k-1} \vee \mathbf{X}_k)|\mathcal{G}_{k-1}] \\
= \ & E[\Lambda_k^{\boldsymbol{\delta}}\Lambda_{k-1}^{\boldsymbol{\theta}}|\mathcal{G}_{k-1}] \\
= \ & \Lambda_{k-1}^{\boldsymbol{\theta}}\Lambda_{k-1}^{\boldsymbol{\delta}} = \Lambda_{k-1}^{\boldsymbol{\theta},\boldsymbol{\delta}} .
\end{aligned}$$

□

For any pair such $(\boldsymbol{\theta}, \boldsymbol{\delta})$ of $(\mathcal{G}, P)$-adapted processes, we can define a new probability measure $Q \sim P$ on $\mathcal{G}_k$, for each $k \in \mathcal{T}$, by

$$\left.\frac{dQ}{dP}\right|_{\mathcal{G}_k} := \Lambda_k^{\theta,\boldsymbol{\delta}} . \tag{10}$$

The above construction of Q extends one proposed in Elliott et al. (2005) for option valuation when the return process is governed by a regime-switching Gaussian process, in the sense that we accommodate our method to price two sources of risk, namely, the risks due to the random shocks ε and the transitions of the economic regimes $\mathbf{X}$. Indeed, the market price of factor risk corresponding to the process $\mathbf{Y}_k$ is embedded in the parameter $\boldsymbol{\theta}_k$, while $\boldsymbol{\delta}_k$ represents the market price of the regime shift from $\mathbf{X}_{k-1}$ to $\mathbf{X}_k$.

To value bonds and other interest-rate sensitive securities, we need to consider the risk-neutral dynamics of the interest rate and the economic regimes. The following proposition gives the required result.

Propostion 3.1

(a) *Under the risk-neutralized probability measure Q, the state vector process has the following dynamics:*

$$\mathbf{Y}_{k+1} \sim \boldsymbol{\alpha}(\mathbf{X}_k) + \boldsymbol{\Gamma}(\mathbf{X}_k)\mathbf{Y}_k + \boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\varepsilon}^*_{k+1} \ , \quad k = 0, 1, \cdots, T-1 \ ,$$

where "$\sim$" means equality in distribution; ε^*_k's are independent and identically distributed random variables with $\boldsymbol{\varepsilon}_k \sim N(\mathbf{0}, \mathbf{I})$ under Q and are independent of $\mathbf{X}_k$.

(b) *The elements of the transition matrix $\mathbf{C}_k = (c_{ijk})_{1\le i,j\le N}$ of the Markov chain $\mathbf{X}$ under Q are given by:*

$$c_{ijk} = \frac{a_{ijk}\exp(\delta_{ijk})}{\sum_{j=1}^{N} a_{ijk}\exp(\delta_{ijk})} \ . \tag{11}$$

Proof Firstly, we evaluate the conditional moment generating function of the factor process under Q, denoted as $M^Q_{\mathbf{Y}_{k+1}|\mathcal{G}_k}(\mathbf{z})$. Write E^Q for expectation under Q. By a version of the Bayes' rule, we have:

$$\begin{aligned}
& M^Q_{\mathbf{Y}_{k+1}|\mathcal{G}_k}(\mathbf{z}) \\
:= \ & E^Q[\exp(\mathbf{z}'\mathbf{Y}_{k+1})|\mathcal{G}_k] \\
= \ & \frac{E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\frac{dQ}{dP}|\mathcal{G}_k]}{E[\frac{dQ}{dP}|\mathcal{G}_k]} \\
= \ & \frac{E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\Lambda^{\boldsymbol{\theta},\boldsymbol{\delta}}_{k+1}|\mathcal{G}_k]}{E[\Lambda^{\boldsymbol{\theta},\boldsymbol{\delta}}_{k+1}|\mathcal{G}_k]} \\
= \ & \frac{E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}\lambda^{\boldsymbol{\delta}}_{k+1}\Lambda^{\boldsymbol{\theta},\boldsymbol{\delta}}_{k}|\mathcal{G}_k]}{\Lambda^{\boldsymbol{\theta},\boldsymbol{\delta}}_{k}} \\
= \ & E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}\lambda^{\boldsymbol{\delta}}_{k+1}|\mathcal{G}_k] \\
= \ & E[E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}\lambda^{\boldsymbol{\delta}}_{k+1}|\mathcal{G}_k \vee \sigma\{\boldsymbol{\varepsilon}_{k+1}\}]|\mathcal{G}_k] \ .
\end{aligned}$$

Since $\lambda^{\boldsymbol{\theta}}_{k+1}$ depends only on $\mathbf{X}$ up to time k and $\boldsymbol{\varepsilon}$ is independent of $\mathbf{X}$ under $\mathcal{P}$,

$$\begin{aligned}
& E[E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}\lambda^{\boldsymbol{\delta}}_{t+1}|\mathcal{G}_k \vee \sigma\{\boldsymbol{\varepsilon}_{k+1}\}]|\mathcal{G}_k] \\
= \ & E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}E[\lambda^{\boldsymbol{\delta}}_{k+1}|\mathcal{G}_k \vee \sigma\{\boldsymbol{\varepsilon}_{k+1}\}]|\mathcal{G}_k] \\
= \ & E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}|\mathcal{G}_k] \ .
\end{aligned}$$

The last equality is due to the fact that $E[\lambda^{\boldsymbol{\delta}}_{k+1}|\mathcal{G}_k \vee \sigma\{\varepsilon_{k+1}\}] = 1$ under the assumption that ε is independent of $\mathbf{X}$ under $\mathcal{P}$.

Consequently,

$$
\begin{aligned}
& M^Q_{\mathbf{Y}_{k+1}|\mathcal{G}_k}(\mathbf{z}) \\
= \; & E[\exp(\mathbf{z}'\mathbf{Y}_{k+1})\lambda^{\boldsymbol{\theta}}_{k+1}|\mathcal{G}_k] \\
= \; & E\left[\exp\left(\mathbf{z}'\mathbf{Y}_{k+1} - \frac{1}{2}\boldsymbol{\theta}'_k(\mathbf{X}_k)\boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k)\boldsymbol{\theta}_k(\mathbf{X}_k) - \boldsymbol{\theta}'_k(\mathbf{X}_k)\boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\varepsilon}_{k+1}\right)\Big|\mathcal{G}_k\right] \\
= \; & \exp\left[\mathbf{z}'\mathbf{m}_k(\mathbf{X}_k) - \frac{1}{2}\boldsymbol{\theta}'_k(\mathbf{X}_k)\boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k)\boldsymbol{\theta}_k(\mathbf{X}_k)\right] M_{\boldsymbol{\varepsilon}_{k+1}}\Big(\boldsymbol{\Sigma}'(\mathbf{X}_k)(\mathbf{z} - \boldsymbol{\theta}_k(\mathbf{X}_k))\Big) \\
= \; & \exp\left[\mathbf{z}'\Big(\boldsymbol{\alpha}(\mathbf{X}_k) + \boldsymbol{\Gamma}(\mathbf{X}_k)\mathbf{Y}_k\Big) + \frac{1}{2}\mathbf{z}'\boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k)\mathbf{z}\right] .
\end{aligned}
$$

This is the moment generating function for a multivariate Gaussian random variable with mean vector $\boldsymbol{\alpha}(\mathbf{X}_k) + \boldsymbol{\Gamma}(\mathbf{X}_k)\mathbf{Y}_k$ and covariance matrix $\boldsymbol{\Sigma}(\mathbf{X}_k)\boldsymbol{\Sigma}'(\mathbf{X}_k)$. This proves (a).

Given that the Markov chain $\mathbf{X}$ is in the i^{th} regime at time k, the conditional moment generating function of $\mathbf{X}_{k+1}$ under P evaluated at $\boldsymbol{\delta}_k$ is given by:

$$
\begin{aligned}
& M_{\mathbf{X}_{k+1}|\mathcal{G}_k \vee \sigma\{\mathbf{X}_k\}}(\boldsymbol{\delta}_k)|_{\mathbf{X}_k=\mathbf{e}_i} \\
= \; & E[\exp(\langle\boldsymbol{\delta}_k, \mathbf{X}_{k+1}\rangle)|\mathcal{G}_k \vee \sigma\{\mathbf{X}_k\}]|_{\mathbf{X}_k=\mathbf{e}_i} \\
= \; & E[\mathbf{X}'_k\mathbf{L}'_k\mathbf{X}_{k+1}|\mathcal{G}_k \vee \sigma\{\mathbf{X}_k\}]|_{\mathbf{X}_k=\mathbf{e}_i} \\
= \; & \sum_{j=1}^{N} a_{ijk}(\mathbf{e}'_i\mathbf{L}'_k\mathbf{e}_j) \\
= \; & \sum_{j=1}^{N} a_{ijk}\exp(\delta_{ijk}) .
\end{aligned}
$$

Therefore, by a version of the Bayes' rule again, under Q the transition probabilities of the Markov chain $\mathbf{X}$ are:

$$\begin{aligned}
c_{ijk} &= E^Q[I_{\{\mathbf{X}_{k+1}=\mathbf{e}_j\}}|\mathcal{G}_k \vee \sigma\{\mathbf{X}_k\}]|_{\mathbf{X}_k=\mathbf{e}_i} \\
&= \frac{E[I_{\{\mathbf{X}_{k+1}=\mathbf{e}_j\}} \exp(\langle \boldsymbol{\delta}_k, \mathbf{X}_{k+1}\rangle)|\mathcal{G}_k \vee \sigma\{\mathbf{X}_k\}]|_{\mathbf{X}_k=\mathbf{e}_i}}{M_{\mathbf{X}_{k+1}|\mathcal{G}_k\vee\sigma\{\mathbf{X}_k\}}(\boldsymbol{\delta}_k)|_{\mathbf{X}_k=\mathbf{e}_i}} \\
&= \frac{E[I_{\{\mathbf{X}_{k+1}=\mathbf{e}_j\}}\mathbf{X}'_k\mathbf{L}'_k\mathbf{X}_{k+1}|\mathcal{G}_k \vee \sigma\{\mathbf{X}_k\}]|_{\mathbf{X}_k=\mathbf{e}_i}}{\sum_{j=1}^N a_{ijk}\exp(\delta_{ijk})} \\
&= \frac{a_{ijk}\mathbf{e}'_i\mathbf{L}'_k\mathbf{e}_j}{\sum_{j=1}^N a_{ijk}\exp(\delta_{ijk})} \\
&= \frac{a_{ijk}\exp(\delta_{ijk})}{\sum_{j=1}^N a_{ijk}\exp(\delta_{ijk})} \ .
\end{aligned}$$

Hence (b) follows. $\square$

Thus, we have showed that under the risk neutral measure given by a double Esscher transform the state process $\mathbf{Y}_k$ has a conditionally multivariate Gaussian distribution with the same covariance matrix as under the physical measure P and a shifted mean vector. Moreover, we note that after the change of measure the market price of risk factor $\boldsymbol{\theta}_k(\mathbf{X}_k)$ has been eliminated from the state process dynamics, while the price of regime switching risk is present in the risk neutral, time-dependent, transition matrix $\mathbf{C}_k$.

Let $B(k,T)$ be the time-k price of a zero-coupon bond with maturity at time T. Write $Z(k,n)$ for the discounted bond price at time k; that is,

$$Z(k,T) := \exp\left(-\sum_{t=0}^{k-1} r_t\right)B(k,T) \ .$$

If Q is a new probability measure under which the discounted bond price process $Z(T) := \{Z(k,T)\}_{k\in\mathcal{T}}$ is a martingale with respect to the filtration $\mathcal{G}$, then we have:

$$B(k,T) = E^Q\left[\exp\left(-\sum_{t=k}^{T-1} r_t\right)|\mathcal{G}_k\right] \ . \tag{12}$$

Recall that E^Q is the expectation under Q. We assume that $B(T,T) = 1$ and $B(T-1,T) = \exp(-r_{T-1})$. In the next section, we shall derive a formula for $B(k,T)$.

4 Exponential Affine-Form Bond Prices

Several continuous time models give rise to explicit exponential-affine forms for the bond prices (see, for example, Elliott and van der Hoek (2001), and the references therein). This led authors to hypothesize an exponential affine form for bond prices which will certainly not be the case for all short rate dynamics. Dai *et al.* (2007) and Ang *et al.* (2007) suppose that the price of a zero-coupon bond is of an exponential form, where the free coefficient is regime dependent while the short rate coefficient is regime-invariant. Based on the martingale property of discounted bond prices under a risk neutral measure, they derived some recursive relationships for these coefficients. However, their approach cannot be implemented in our setting since the state-dependent component of the mean under a price kernel selected by the double Esscher transform, namely, $\mathbf{\Gamma}(\mathbf{X}_k)\mathbf{Y}_k$, is regime-dependent and the risk-neutral transition matrix is time-dependent. Moreover, if we follow the traditional approach in our modeling framework, we need to deal with the situation that both coefficients are subject to regime shifts. However, in this case there will be no analytic recursive relationships between the coefficients of the exponential affine form of the bond price. So one may need some approximation procedures to the bond price. For example, Bansal and Zhou (2002) obtained closed-form expressions using a log-linear approximation.

Our approach is different from the approach adopted in the existing literature as we derive an exponential affine form for the conditional bond price given the future sample path of the Markov chain. We then use this exponential affine form to obtain the bond price $B(k, n)$ as a weighted average of exponential affine forms, where the coefficients of the exponential affine forms satisfy some recursion relations and the probability weights are given by the transition probabilities of the Markov chain.

Firstly, we define the conditional bond price given the future sample path of the Markov chain.

For each $k = 0, 1, \cdots, T$, let $\mathcal{H}_k := \mathcal{F}_T^{\mathbf{X}} \vee \mathcal{F}_k^{\mathbf{Y}}$, which is the enlarged information set generated by histories of the Markov chain up to the maturity time T and the short rate process up to time k. Then the conditional price of the zero-coupon bond at time k given the enlarged information set $\mathcal{H}_k$, denoted as $\tilde{B}(k,T)$, is defined by setting

$$\tilde{B}(k,T) := E^Q\Bigg[\exp\bigg(-\sum_{t=k}^{T-1} r_t\bigg)|\mathcal{H}_k\Bigg]\ , \quad k=0,1,\cdots,T\ . \tag{13}$$

The following proposition gives an exponential affine form for the conditional bond price $\tilde{B}(k,T)$.

Propostion 4.1 *The conditional bond price $\tilde{B}(k,T)$ has the following exponential affine form:*

$$\tilde{B}(k,T) = \exp\Big(\kappa_1(\mathbf{X}_k) + \boldsymbol{\kappa}_2'(\mathbf{X}_k)\mathbf{Y}_k\Big)\ , \quad k=0,1,\cdots,T\ , \tag{14}$$

where the stochastic coefficients $\{\kappa_1(\mathbf{X}_k)\}_{k=0,1,\cdots,T}$ and $\{\boldsymbol{\kappa}_2(\mathbf{X}_k)\}_{k=0,1,\cdots,T}$ satisfy the following system of coupled stochastic recursions:

$$\begin{aligned}
\kappa_1(\mathbf{X}_{k-1}) &= \kappa_1(\mathbf{X}_k) + \boldsymbol{\kappa}_2'(\mathbf{X}_k)\boldsymbol{\alpha}(\mathbf{X}_{k-1}) \\
&+ \frac{1}{2}\boldsymbol{\kappa}_2'(\mathbf{X}_k)\boldsymbol{\Sigma}(\mathbf{X}_{k-1})\boldsymbol{\Sigma}'(\mathbf{X}_{k-1})\boldsymbol{\kappa}_2(\mathbf{X}_k) - d_0(\mathbf{X}_{k-1})\ , \\
\boldsymbol{\kappa}_2(\mathbf{X}_{k-1}) &= \boldsymbol{\Gamma}'(\mathbf{X}_{k-1})\boldsymbol{\kappa}_2(\mathbf{X}_k) - \mathbf{d}_1(\mathbf{X}_{k-1})\ , \quad k=1,2,\cdots,T\ ,
\end{aligned}$$

with initial conditions $\kappa_1(\mathbf{X}_T) = \boldsymbol{\kappa}_2(\mathbf{X}_T) = 0$.

Proof We prove the result by the backward induction. Since $\tilde{B}(T,T)=1$, the result is obviously true for $k=T$. Suppose the result holds for $k=n$. We wish to prove that it is also true for $k=n-1$. Now by repeated conditional expectation, the assumption for $k=n$ and Proposition 2.1,

$$\begin{aligned}
\tilde{B}(n-1,T) &= E^Q\Bigg[\exp\bigg(-\sum_{t=n-1}^{T-1} r_t\bigg)|\mathcal{H}_{n-1}\Bigg] \\
&= E^Q\Bigg\{\exp(-r_{n-1})E^Q\Bigg[\exp\bigg(-\sum_{t=n}^{T-1} r_t\bigg)|\mathcal{H}_n\Bigg]|\mathcal{H}_{n-1}\Bigg\} \\
&= \exp\Big(-r_{n-1}\Big)E^Q\Big[\exp\Big(\kappa_1(\mathbf{X}_n) + \boldsymbol{\kappa}_2'(\mathbf{X}_n)\mathbf{Y}_n\Big)|\mathcal{H}_{n-1}\Big] \\
&= \exp\Big[\kappa_1(\mathbf{X}_n) + \boldsymbol{\kappa}_2'(\mathbf{X}_n)\boldsymbol{\alpha}(\mathbf{X}_{n-1})\Big] \\
&\times \exp\Big[\frac{1}{2}\boldsymbol{\kappa}_2'(\mathbf{X}_n)\boldsymbol{\Sigma}(\mathbf{X}_{n-1})\boldsymbol{\Sigma}'(\mathbf{X}_{n-1})\boldsymbol{\kappa}_2(\mathbf{X}_n) - d_0(\mathbf{X}_{n-1})\Big] \\
&\times \exp\Big[\boldsymbol{\kappa}_2'(\mathbf{X}_n)\boldsymbol{\Gamma}(\mathbf{X}_{n-1}) - \mathbf{d}_1'(\mathbf{X}_{n-1})\Big] \\
&= \exp\Big(\kappa_1(\mathbf{X}_{n-1}) + \boldsymbol{\kappa}_2'(\mathbf{X}_{n-1})\mathbf{Y}_{n-1}\Big)\ .
\end{aligned}$$

Hence the result follows. $\square$

Note that $\tilde{B}(k,T)$ is a function of $\mathbf{X}_k$, $\mathbf{X}_{k+1}$, $\cdots$, $\mathbf{X}_{T-1}$, written say

$$\tilde{B}(k,T,\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1}) \ ,$$

and that the coefficients $\kappa_1(\mathbf{X}_k)$ and $\boldsymbol{\kappa}_2(\mathbf{X}_k)$, $k=0,1,\cdots,T-1$, are measurable with respect to the tail σ-algebra generated by $\mathbf{X}_k$, $\mathbf{X}_{k+1}$, $\cdots$ and $\mathbf{X}_{T-1}$. So they can be represented as functions of $\mathbf{X}_k$, $\mathbf{X}_{k+1}$, $\cdots$ and $\mathbf{X}_{T-1}$; that is,

$$\begin{aligned} \kappa_1(\mathbf{X}_k) &= \kappa_1(\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1}) \ , \\ \boldsymbol{\kappa}_2(\mathbf{X}_k) &= \boldsymbol{\kappa}_2(\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1}) \ , \quad k=0,1,\cdots,T-1 \ . \end{aligned} \tag{15}$$

Then for each $k=0,1,\cdots,T-1$, given $\mathcal{H}_k=\mathcal{F}_T^{\mathbf{X}}\vee\mathcal{F}_k^r$ the conditional bond price $\tilde{B}(k,T,\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1})$ can be represented as follows:

$$\begin{aligned} &\tilde{B}(k,T,\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1}) \\ = \ &\exp\left(\kappa_1(\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1})+\boldsymbol{\kappa}_2'(\mathbf{X}_k,\mathbf{X}_{k+1},\cdots,\mathbf{X}_{T-1})\mathbf{Y}_k\right) \ . \end{aligned} \tag{16}$$

In practice, we do not anticipate future states of the economy. So to implement our model in practical situations, we need to evaluate the bond price given the information set $\mathcal{G}_k$. We derive the result in the following corollary.

Corollary 4.1 *Given $\mathcal{G}_k$, the bond price has the following form:*

$$\begin{aligned} &B(k,T) \\ = &\sum_{i_k,i_{k+1},\cdots,i_{T-1}=1}^{N}\left[\prod_{l=k}^{T-1}c_{i_l i_{l+1} k}\right]\exp\left(\kappa_1(\mathbf{e}_{i_k},\mathbf{e}_{i_{k+1}},\cdots,\mathbf{e}_{i_{T-1}})+\boldsymbol{\kappa}_2'(\mathbf{e}_{i_k},\mathbf{e}_{i_{k+1}},\cdots,\mathbf{e}_{i_{T-1}})\mathbf{Y}_k\right) \\ = &\sum_{i_k,i_{k+1},\cdots,i_{T-1}=1}^{N}\left[\prod_{l=k}^{T-1}c_{i_l i_{l+1} k}\right]\tilde{B}(k,T,\mathbf{e}_{i_k},\mathbf{e}_{i_{k+1}},\cdots,\mathbf{e}_{i_{T-1}}) \ . \end{aligned} \tag{17}$$

Proof This result is obtained from taking expectation of $\tilde{B}(k,n)$ conditioning on $\mathcal{G}_k$ under $\mathcal{Q}$ and by enumerating transition probabilities of the Markov chain from time k to time $T-1$. $\square$

The corollary states that the bond price given information $\mathcal{G}_k$ up to time k is the weighted average of exponential affine forms with weights as transition probabilities of future paths of the Markov chain. This is a practical form of the bond price and can be computed by counting possible future paths of the chain. However, when testing the empirical performance of our pricing methodology, one can impose some constraints on the structure of the model parameters. For example, we consider some parametrization for the real-world transition probabilities and we further assume the risk neutral probabilities are time independent, one could apply similar estimation techniques as in Dai *et al.* (2007) for estimating the model parameters. However, our pricing result from Corollary 4.1 provides a more flexible result since is was established under a very general discrete time affine term structure process. Our proposed model appears more flexible than the one in Dai and Singleton (2007) since we do not require a time-dependent component of the risk-neutralized conditional mean.

5 Continuous-Time Extension and Pricing Kernel

In this section we consider a continuous-time extension of the proposed discrete-time, multi-factor, interest rate model. Here the time index set $\mathcal{T}$ is $[0, T]$, for $T < \infty$, and again $(\Omega, \mathcal{F}, P)$ is a complete probability space.

Let $\mathbf{X} := \{\mathbf{X}(t)\}_{t \in \mathcal{T}}$ be a continuous-time, finite-state, Markov chain on $(\Omega, \mathcal{F}, P)$ whose states represent different states of the economy. Again we take the state space of $\mathbf{X}$ as $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_N\}$. To specify the probability law of the chain $\mathbf{X}$, we consider a family of rate matrices $\mathbf{A}(t) := [a_{ij}(t)]_{i,j=1,2,\cdots,N}$, $t \in \mathcal{T}$, where $a_{ij}(t)$ is the instantaneous rate matrix of the chain from state $\mathbf{e}_i$ to state $\mathbf{e}_j$ at time t. Then Elliott et al. (1994) gave the following semimartingale dynamics for the chain $\mathbf{X}$:

$$\mathbf{X}(t) = \mathbf{X}(t) + \int_0^t \mathbf{A}(u)\mathbf{X}(u)du + \mathbf{K}(t) \ . \tag{18}$$

Here $\mathbf{K} := \{\mathbf{K}(t)\}_{t \in \mathcal{T}}$ is an $\Re^N$-valued, $(\mathcal{F}^{\mathbf{X}}, P)$-martingale, where $\mathcal{F}^{\mathbf{X}}$ is the right-continuous, P-completed, filtration generated by the chain $\mathbf{X}$.

Let $\{\mathbf{W}(t)\}_{t \in \mathcal{T}}$ be an $\Re^L$-valued Brownian motion on $(\Omega, \mathcal{F}, P)$. Then we assume that under P, a multivariate factor process $\{\mathbf{Z}(t)\}_{t \in \mathcal{T}}$ is governed by the following multivariate, mean-reverting, Markov, regime-switching, diffusion process:

$$d\mathbf{Z}(t) = [\boldsymbol{\alpha}(\mathbf{X}(t)) + \boldsymbol{\Sigma}'(\mathbf{X}(t))\boldsymbol{\theta}(\mathbf{X}(t)) + \boldsymbol{\Gamma}(\mathbf{X}(t))\mathbf{Z}(t)]dt + \boldsymbol{\Sigma}'(\mathbf{X}(t))d\mathbf{W}(t) \ ,$$

where $\boldsymbol{\alpha}(\mathbf{X}(t))$, $\boldsymbol{\Sigma}(\mathbf{X}(t))$, $\boldsymbol{\theta}(\mathbf{X}(t))$ and $\boldsymbol{\Gamma}(\mathbf{X}(t))$ are defined as in Section 2, except that the time index t here takes a value in the time index set $\mathcal{T}$ of the continuous-time economy.

Write $\mathcal{F}^{\mathbf{Z}}$ for the right-continuous, P-completed, filtration $\{\mathcal{F}^{\mathbf{Z}}(t)\}_{t\in\mathcal{T}}$ generated by the process of the market prices of factors. For each $t \in \mathcal{T}$, write $\mathcal{H}(t) := \mathcal{F}^{\mathbf{Z}}(t) \vee \mathcal{F}^{\mathbf{X}}(t)$ and $\mathcal{H} := \{\mathcal{H}(t)\}_{t\in\mathcal{T}}$. In the continuous-time case, we assume that the process of market prices of factors $\{\boldsymbol{\theta}(\mathbf{X}(t))\}_{t\in\mathcal{T}}$ are $\mathcal{H}$-predictable.

As in Section 2, we suppose that the short-term interest rate process $\{r(t)\}_{t\in\mathcal{T}}$ has the following affine form in terms of the market factors $\mathbf{Z}$:

$$r(t) := d_0 + \mathbf{d}_1'\mathbf{Z}(t) \ ,$$

where $d_0 \in \Re$ and $\mathbf{d}_1 \in \Re^L$.

In what follows, we specify a pricing kernel by the product of two density processes, one for a measure change of the Brownian motion $\{\mathbf{W}(t)\}_{t\in\mathcal{T}}$ and one for a measure change for the Markov chain $\mathbf{X}$. The pricing kernel takes into account explicitly both the diffusion risk and regime-switching risk.

Define an $\mathcal{H}$-adapted process $\Lambda^{\boldsymbol{\theta}} := \{\Lambda^{\boldsymbol{\theta}}(t)\}_{t\in\mathcal{T}}$ associated with $\boldsymbol{\theta}$ by putting

$$\Lambda^{\boldsymbol{\theta}}(t) := \exp\left(- \int_0^t \boldsymbol{\theta}'(s)d\mathbf{W}(s) - \frac{1}{2}\int_0^t \boldsymbol{\theta}^2(s)ds \right) \ . \tag{19}$$

From its definition, it is not difficult to see that the factor price of risk process $\{\boldsymbol{\theta}(t)\}_{t\in\mathcal{T}}$ satisfies the Novikov condition. Consequently, $\Lambda^{\boldsymbol{\theta}}$ is an $(\mathcal{H}, P)$-martingale, and

$$\mathrm{E}[\Lambda^{\boldsymbol{\theta}}(T)] = 1 \ .$$

We now define a density process for a measure change for the Markov chain $\mathbf{X}$. For each $i, j = 1, 2, \cdots, N$, we consider a real-valued, $\mathcal{F}^{\mathbf{Z}}$-predictable, bounded process $\{c_{ij}(t)\}_{t\in\mathcal{T}}$ on $(\Omega, \mathcal{F}, P)$ such that for each $t \in \mathcal{T}$,

1. $c_{ij}(t) \geq 0$, for $i \neq j$;

2. $\sum_{j=1}^N c_{ij}(t) = 0$, so $c_{ii}(t) \leq 0$.

Then $\mathbf{C}(t) := \{c_{ij}(t)\}_{i,j=1,2,\cdots,N}$, $t \in \mathcal{T}$, is a second family of rate matrices of the chain $\mathbf{X}$. We wish to introduce a new probability measure under which $\mathbf{C}(t)$, $t \in \mathcal{T}$, are a family of rate matrices of the chain $\mathbf{X}$. The development here follows that of Dufour and Elliott (1999), where a version of Girsanov's transform for the Markov chain was adopted.

Define, for each $t \in \mathcal{T}$, the following matrix:

$$\begin{aligned} \mathbf{D}^{\mathbf{C}}(t) &:= [c_{ij}(t)/a_{ij}(t)]_{i,j=1,2,\ldots,N} \\ &= [d^{\mathbf{C}}_{ij}(t)] \quad \text{say} . \end{aligned}$$

Note that $a_{ij}(t) > 0$, for each $t \in \mathcal{T}$, so $\mathbf{D}(t)$ is well-defined.

For each $t \in \mathcal{T}$, let

$$\mathbf{d}^{\mathbf{C}}(t) := (d^{\mathbf{C}}_{11}(t), d^{\mathbf{C}}_{22}(t), \cdots, d^{\mathbf{C}}_{NN}(t))' \in \Re^N .$$

Write, for each $t \in \mathcal{T}$,

$$\mathbf{D}^{\mathbf{C}}_0(t) := \mathbf{D}^{\mathbf{C}}(t) - \mathbf{diag}(\mathbf{d}^{\mathbf{C}}(t)) ,$$

where $\mathbf{diag}(\mathbf{y})$ is a diagonal matrix with diagonal elements given by the vector $\mathbf{y}$.

Consider the vector-valued counting process, $\{\mathbf{N}(t)\}_{t\in\mathcal{T}}$, on $(\Omega, \mathcal{F}, P)$, where for each $t \in \mathcal{T}$, $\mathbf{N}(t) := (N_1(t), N_2(t), \cdots, N_N(t))' \in \Re^N$ and $N_j(t)$ counts the number of jumps of the chain $\mathbf{X}$ to state $\mathbf{e}_j$ up to time t, for each $j = 1, 2, \cdots, N$. Then

$$\mathbf{N}(t) = \int_0^t (\mathbf{I} - \mathbf{diag}(\mathbf{X}(u-)))' d\mathbf{X}(u) . \tag{20}$$

Here $\mathbf{N}(0) = \mathbf{0} \in \Re^N$.

The following lemma was due to Dufour and Elliott (1999). We state the result here without giving the proof.

Lemma 5.1 *Let* $\mathbf{A}_0(t) := \mathbf{A}(t) - \mathbf{diag}(\mathbf{a}(t))$, *where* $\mathbf{a}(t) := (a_{11}(t), a_{22}(t), \cdots, a_{NN}(t))' \in \Re^N$, *for each* $t \in \mathcal{T}$. *Then the process* $\{\tilde{\mathbf{N}}(t)\}_{t\in\mathcal{T}}$ *defined by putting*

$$\tilde{\mathbf{N}}(t) := \mathbf{N}(t) - \int_0^t \mathbf{A}_0(u)\mathbf{X}(u-)du , \quad t \in \mathcal{T} , \tag{21}$$

is an $\Re^N$*-valued,* $(\mathcal{F}^{\mathbf{X}}, P)$*-martingale.*

Consider the $\mathcal{F}^{\mathbf{X}}$-adapted process $\{\Lambda^{\mathbf{C}}(t)\}_{t\in\mathcal{T}}$ on $(\Omega, \mathcal{F}, P)$ associated with $\{\mathbf{C}(t)\}_{t\in\mathcal{T}}$ defined by setting

$$\Lambda^{\mathbf{C}}(t) = 1 + \int_0^t \Lambda^{\mathbf{C}}(u-)[\mathbf{D}_0^{\mathbf{C}}(u)\mathbf{X}(u-) - \mathbf{1}]' d\tilde{\mathbf{N}}(u) \ .$$

Here $\mathbf{1} := (1, 1, \ldots, 1)' \in \Re^N$.

Then the following result is an immediate consequence of Lemma 5.1 and the boundedness of $c_{ij}(t)$, for each $i, j = 1, 2, \cdots, N$ and $t \in \mathcal{T}$.

Lemma 5.2 $\Lambda^{\mathbf{C}}$ *is an* $(\mathcal{F}^{\mathbf{X}}, P)$*-martingale.*

Here $\Lambda^{\mathbf{C}}$ is used as a density process for a measure change for the chain $\mathbf{X}$.

Consider the $\mathcal{H}$-adapted process $\{\Lambda^{\boldsymbol{\theta},\mathbf{C}}(t)\}_{t\in\mathcal{T}}$ on $(\Omega, \mathcal{F}, P)$ defined by:

$$\Lambda^{\boldsymbol{\theta},\mathbf{C}}(t) := \Lambda^{\boldsymbol{\theta}}(t) \cdot \Lambda^{\mathbf{C}}(t) \ , \quad t \in \mathcal{T} \ .$$

Our assumptions ensure that $\Lambda^{\boldsymbol{\theta},\mathbf{C}}$ is an $(\mathcal{H}, P)$-martingale.

We now define a probability measure $Q^{\boldsymbol{\theta},\mathbf{C}}$ absolutely continuous with respect to P on $\mathcal{H}(T)$ as:

$$\left.\frac{dQ^{\boldsymbol{\theta},\mathbf{C}}}{dP}\right|_{\mathcal{H}(T)} := \Lambda^{\boldsymbol{\theta},\mathbf{C}}(T) \ . \tag{22}$$

This is a density process for a measure change for both the Brownian motion $\mathbf{W}$ and the Markov chain $\mathbf{X}$.

The following theorem gives the probability laws of the Brownian motion $\mathbf{W}$ and the chain $\mathbf{X}$ under the new measure $Q^{\boldsymbol{\theta},\mathbf{C}}$.

Theorem 5.1 *The process defined by*

$$\mathbf{W}^{\boldsymbol{\theta}}(t) := \mathbf{W}(t) + \int_0^t \boldsymbol{\theta}(u) du \ , \quad t \in \mathcal{T} \ ,$$

is an $\Re^L$*-valued,* $(\mathcal{H}, Q^{\boldsymbol{\theta},\mathbf{C}})$*-Brownian motion. Under* $Q^{\boldsymbol{\theta},\mathbf{C}}$*, the chain* $\mathbf{X}$ *has a family of rate matrices* $\mathbf{C}(t)$*,* $t \in \mathcal{T}$*, and can be represented as:*

$$\mathbf{X}(t) = \mathbf{X}(0) + \int_0^t \mathbf{C}(u)\mathbf{X}(u-) du + \mathbf{M}^{\mathbf{C}}(t) \ ,$$

where $\{\mathbf{M}^{\mathbf{C}}(t)\}_{t\in\mathcal{T}}$ *is an* $\Re^N$*-valued,* $(\mathcal{H}, Q^{\boldsymbol{\theta},\mathbf{C}})$*-martingale.*

Proof The results follow directly from Girsanov's theorem for the Brownian motion and a Girsanov transform for the Markov chain. $\square$

Corollary 5.1 *Under $Q^{\boldsymbol{\theta},\mathbf{C}}$, the factor price of risk process evolves over time as:*

$$d\mathbf{Z}(t) = [\boldsymbol{\alpha}(\mathbf{X}(t)) + \boldsymbol{\Gamma}(\mathbf{X}(t))\mathbf{Z}(t)]dt + \boldsymbol{\Sigma}(\mathbf{X}(t))d\mathbf{W}^{\boldsymbol{\theta}}(t) \ .$$

Proof The result follows from Theorem 5.1. $\square$

6 Stochastic Flows and Exponential Affine Bond Prices

In this section, we adopt the concept of stochastic flows to derive an exponential affine formula for the bond price. The results developed here are extensions of those in Elliott and van der Hoek (2001) and Elliott and Siu (2009).

Firstly, we recall from the last section that under $Q^{\boldsymbol{\theta},\mathbf{C}}$,

$$d\mathbf{Z}(t) = [\boldsymbol{\alpha}(\mathbf{X}(t)) + \boldsymbol{\Gamma}(\mathbf{X}(t))\mathbf{Z}(t)]dt + \boldsymbol{\Sigma}(\mathbf{X}(t))d\mathbf{W}^{\boldsymbol{\theta}}(t) \ ,$$
$$d\mathbf{X}(t) = \mathbf{C}(t)\mathbf{X}(t)dt + d\mathbf{M}^{\mathbf{C}}(t) \ .$$

Here to develop the exponential affine formula for the bond price, we have to assume that $\boldsymbol{\Gamma}(t) = \boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma}$ is a constant matrix in $\Re^L \otimes \Re^L$. With this assumption, the factor price of risk process under $Q^{\boldsymbol{\theta},\mathbf{C}}$ satisfies the following $\Re^L$-valued, stochastic differential equation:

$$d\mathbf{Z}(t) = [\boldsymbol{\alpha}(\mathbf{X}(t)) + \boldsymbol{\Gamma}\mathbf{Z}(t)]dt + \boldsymbol{\Sigma}(\mathbf{X}(t))d\mathbf{W}^{\boldsymbol{\theta}}(t) \ . \tag{23}$$

Suppose $\mathbf{Z}_{s,t}(\mathbf{z})$, $t, s \in \mathcal{T}$ with $t \geq s$, is a version of the process satisfying the stochastic differential equation (23) with initial condition $\mathbf{Z}_{s,s}(\mathbf{z}) = \mathbf{z}$. Then

$$\mathbf{Z}_{s,t}(\mathbf{z}) = \mathbf{z} + \int_s^t [\boldsymbol{\alpha}(\mathbf{X}(u)) + \boldsymbol{\Gamma}\mathbf{Z}_{s,u}(\mathbf{z})]du + \int_s^t \boldsymbol{\Sigma}(\mathbf{X}(u))d\mathbf{W}^{\boldsymbol{\theta}}(u) \ . \tag{24}$$

Denote the Jacobian of the map $\mathbf{z} \to \mathbf{Z}_{s,t}(\mathbf{z})$ by:

$$\mathbf{D}_{s,t} := \frac{\partial \mathbf{Z}_{s,t}(\mathbf{z})}{\partial \mathbf{z}} \ ,$$

where $\mathbf{D}_{s,s} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix in $\Re^L \otimes \Re^L$.

Differentiating both sides of (24) then gives:

$$\mathbf{D}_{s,t} = \mathbf{I} + \mathbf{\Gamma} \int_s^t \mathbf{D}_{s,u} du \ .$$

Consequently,

$$\mathbf{D}_{s,t} = e^{\mathbf{\Gamma}(t-s)} \ .$$

A price of the zero-coupon bond at time t is determined as:

$$P(t,T) = \mathrm{E}^{\boldsymbol{\theta},\mathbf{C}}\Big[\exp\Big(-\int_t^T r(\mathbf{Z}_{t,u}(\mathbf{z}))du\Big)|\mathcal{H}(t)\Big] \ . \tag{25}$$

Here $\mathrm{E}^{\boldsymbol{\theta},\mathbf{C}}$ is expectation with respect to $Q^{\boldsymbol{\theta},\mathbf{C}}$.

Since $\{(\mathbf{Z}(t), \mathbf{X}(t))\}_{t\in\mathcal{T}}$ is jointly Markov with respect to $\mathcal{H}$,

$$\begin{aligned} P(t,T,\mathbf{z},\mathbf{x}) &:= \mathrm{E}^{\boldsymbol{\theta},\mathbf{C}}\Big[\exp\Big(-\int_t^T r(\mathbf{Z}_{t,u}(\mathbf{z}))du\Big)|\mathbf{Z}(t)=\mathbf{z},\mathbf{X}(t)=\mathbf{x}\Big] \\ &= \mathrm{E}^{\boldsymbol{\theta},\mathbf{C}}\Big[\exp\Big(-\int_t^T r(\mathbf{Z}_{t,u}(\mathbf{z}))du\Big)|\mathcal{H}(t)\Big] \ . \end{aligned} \tag{26}$$

Thanks to the boundedness of the exponential on a finite time interval $[0,T]$, we have:

$$\frac{\partial P}{\partial \mathbf{z}}(t,T,\mathbf{z},\mathbf{x}) = \mathrm{E}^{\boldsymbol{\theta},\mathbf{C}}\Big[\Big(-\mathbf{d}_1'\cdot\int_t^T \mathbf{D}_{t,u}du\Big)\exp\Big(-\int_t^T r(\mathbf{Z}_{t,u}(\mathbf{z}))du\Big)|\mathcal{H}(t)\Big] \ .$$

Write, for each $t\in\mathcal{T}$,

$$\mathbf{A}_2(t,T) := \mathbf{d}_1'\cdot\int_t^T \mathbf{D}_{t,u}du = \mathbf{d}_1'\mathbf{\Gamma}^{-1}[e^{\mathbf{\Gamma}(T-t)} - \mathbf{I}] \in \Re^L \ .$$

Then

$$\frac{\partial P}{\partial \mathbf{z}}(t,T,\mathbf{z},\mathbf{x}) = -\mathbf{A}_2(t,T)P(t,T,\mathbf{z},\mathbf{x}) \ . \tag{27}$$

Let $A_1(t,T,\mathbf{x}) := e^{\tilde{A}_1(t,T,\mathbf{x})}$, for each $(t,\mathbf{x})\in\mathcal{T}\times\mathcal{E}$. Then integrating (27) in $\mathbf{z}$ gives:

$$P(t,T,\mathbf{z},\mathbf{x}) = \exp(A_1(t,T,\mathbf{x}) - \mathbf{A}_2(t,T)\mathbf{z}) \ . \tag{28}$$

To determine an exponential affine form of the bond price $P(t,T,\mathbf{z},\mathbf{x})$, it remains to determine $A_1(t,T,\mathbf{x})$.

Write, for each $i = 1, 2, \cdots, N$, $P_i := P(t, T, \mathbf{z}, \mathbf{e}_i)$ and $\mathbf{P} := (P_1, P_2, \cdots, P_N)' \in \Re^N$. Suppose, for each $i = 1, 2, \cdots, N$,

$$\mathbf{\Sigma}(\mathbf{e}_i) = \mathbf{\Sigma}_i = [\sigma_{kli}]_{k,l=1,2,\cdots,L} \ .$$

Following the arguments in Elliott and van der Hoek (2001) and Elliott and Siu (2009), it is not difficult to check that P_i, $i = 1, 2, \cdots, N$, satisfy the following system of coupled partial differential equations:

$$\frac{\partial P_i}{\partial t} - rP_i + \frac{\partial P_i}{\partial \mathbf{z}}(\boldsymbol{\alpha}_i + \mathbf{\Gamma}\mathbf{z}) + \frac{1}{2}\sum_{k,l=1}^{L}\frac{\partial^2 P_i}{\partial z_k \partial z_l}\sum_{j=1}^{L}\sigma_{kji}\sigma_{lji} + \langle \mathbf{P}, \mathbf{C}(t)\mathbf{e}_i\rangle = 0 \ , \tag{29}$$

with terminal condition $P(T, T, \mathbf{z}, \mathbf{e}_i) = 1$, $i = 1, 2, \cdots, N$.

Write, for each $i = 1, 2, \cdots, N$, $A_{1i} := A_1(t, T, \mathbf{e}_i)$ and $\mathbf{A}_1 := (A_{11}, A_{12}, \cdots, A_{1N})' \in \Re^N$. Suppose, for each $i = 1, 2, \cdots, N$, $J_i := J(t, T, \mathbf{e}_i) = \exp(A_1(T, t, \mathbf{e}_i))$, and $\mathbf{J} := (J_1, J_2, \cdots, J_N)$. Write $\mathbf{A}_2(t, T) := (A_{21}(t, T), A_{22}(t, T), \cdots, A_{2L}(t, T))' \in \Re^L$. Then by (28) and (29), A_{1i}, $i = 1, 2, \cdots, N$, must satisfy the following system of N coupled O.D.E.s:

$$\frac{dA_{1i}}{dt} - d_0 - \mathbf{A}_2'(t, T)\boldsymbol{\alpha}_i + \frac{1}{2}\sum_{k,l=1}^{L} A_{2k}(t, T)A_{2l}(t, T)\sum_{j=1}^{L}\sigma_{kji}\sigma_{lji} + e^{-A_{1i}}\langle \mathbf{J}, \mathbf{C}(t)\mathbf{e}_i\rangle = 0 \ , \tag{30}$$

with terminal conditions:

$$A_1(T, T, \mathbf{e}_i) = 0 \ , \quad i = 1, 2, \cdots, N \ .$$

Write, for each $i = 1, 2, \cdots, N$,

$$f_i(t) := d_0 + \mathbf{A}_2'(t, T)\boldsymbol{\alpha}_i - \frac{1}{2}\sum_{k,l=1}^{L} A_{2k}(t, T)A_{2l}(t, T)\sum_{j=1}^{L}\sigma_{kji}\sigma_{lji} \ ,$$

and consider the following diagonal matrix:

$$\mathbf{diag}(\mathbf{f}(t)) := \mathbf{diag}(f_1(t), f_2(t), \cdots, f_N(t)) \ .$$

Suppose

$$\mathbf{\Delta}(t) := \mathbf{diag}(\mathbf{f}(t)) - \mathbf{C}'(t) \ ,$$

where $\mathbf{C}'(t)$ is the transpose of the matrix $\mathbf{C}(t)$.

Then it can be shown that $\mathbf{J}$ satisfies the following matrix-valued equation:

$$\frac{d\mathbf{J}}{dt} = \mathbf{\Delta}(t)\mathbf{J} \ , \quad \mathbf{J}(0) = \mathbf{1} = (1, 1, \cdots, 1)' \in \Re^N \ .$$

Suppose $\mathbf{\Phi}(t)$ is the fundamental matrix solution of the matrix-valued equation:

$$\frac{d\mathbf{\Phi}(t)}{dt} = \mathbf{\Delta}(t)\mathbf{\Phi}(t) \ , \quad \mathbf{\Phi}(0) = \mathbf{I} \ .$$

Here $\mathbf{I}$ is the $(N \times N)$ identity matrix.

Consequently,

$$\mathbf{J}(t) = \mathbf{\Phi}(t)\mathbf{J}(0) = \mathbf{\Phi}(t)\mathbf{1} \ .$$

Therefore,

$$A_1(t, T, \mathbf{x}) = \sum_{i=1}^{N} \ln \Big(\langle \mathbf{\Phi}(t)\mathbf{1}, \mathbf{e}_i \rangle \Big) \langle \mathbf{X}(t), \mathbf{e}_i \rangle \ .$$

7 Summary

We developed a discrete-time, Markov, regime-switching, affine term structure model for valuing bonds and other interest-rate sensitive securities. The model parameters were modulated by a discrete-time, finite-state, Markov chain whose states represent different states of an economy. The proposed model incorporates the impact of structural changes in (macro)-economic conditions on the interest rate dynamics and includes some econometric advantages. We introduced a double Esscher transform to determine a price kernel for valuation. The double Esscher transform was defined by the product of two density processes for measure changes, one for the interest rate process and another for the Markov chain. Therefore, the double Esscher transform takes into account both the market risk and the long-term economic risk. Using backward induction, we provided a simple derivation of a weighted average of exponential-affine forms for zero coupon bond prices which could be used for empirical investigations. We also discussed a continuous-time version of the proposed model and the specification of the pricing kernel using the product of two density processes. An exponential affine bond pricing formula was derived using the concepts of stochastic flows.

The current paper focused on the theoretical development of regime-switching term structure models and exponential-affine forms of bond prices in both discrete-time and continuous-time setting. For practical implementation of the theoretical results developed in this paper, one important issue is how to estimate the parameters in the model. In particular, one may develop some practically useful methods, based on sound econometric and statistical theories, to estimate the model parameters and investigate the statistical properties, such as the asymptotic properties, of the estimates. These represent potential topics for future research.

Acknowledgement: The authors would like to thank the referees for their helpful comments. Robert J. Elliott and Tak Kuen Siu wish to acknowledge the Discovery Grant from the Australian Research Council (ARC), (Project No.: DP1096243). Alex Badescu wishes to thank NSERC for financial support.

8 References

1. Ang, A., Bekaert, G. 2002a. Short rate nonlinearities and regime switches. Journal of Economic Dynamics and Control 26(7-8), 1243-1274.

2. Ang, A., Bekaert, G. 2002b. International asset allocation with regime shifts. Review of Financial Studies 15(4), 1137-1187.

3. Ang, A., Bekaert, G. 2002c. Regime switches in interest rates. Journal of Business and Economic Statistics 20(2), 163-182.

4. Ang, A., Bekaert, G., Wei, M. 2007. The term structure of real rates and expected inflation. Preprint. Columbia University and NBER.

5. Bansal, R., Zhou, H. 2002b. Term structure of interest rates with regime shifts. Journal of Finance 57(4), 1997-2043.

6. Brennan, M. J., Schwartz, E.S. 1979. A continuous time approach to the pricing of bonds. Journal of Banking and Finance 3, 133-155.

7. Bühlmann, H. 1980. An economic premium principle. ASTIN Bulletin 11(1), 52-60.

8. Bühlmann, H. 1984. The general economic premium principle. ASTIN Bulletin 14(1), 13-21.

9. Bühlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A.N. 1996. No-arbitrage, change of measures and conditional Esscher transforms. CWI Quarterly 9(4), 291-317.

10. Bühlmann, H., Delbaen, F., Embrechts, P., Shiryaev, A.N. 1998. On Esscher transforms in discrete finance models. ASTIN Bulletin 28(2), 171-186.

11. Chapman, D., Pearson, N. 2001. Recent advances in estimating term-structure models. Financial Analysts Journal (July/August), 77-95.

12. Chen, L. 1996. Stochastic Mean and Stochastic Volatility - A Three-Factor Model of the Term Structure of Interest Rates and Its Application to the Pricing of Interest Rate Derivatives. Blackwell Publishers, United Kingdom.

13. Chan, T. 1999. Pricing contingent claims on stocks driven by Lévy processes. Annals of Applied Probability 9(2), 504-528.

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

15. Dai, Q., Singleton, K.J. 2000. Specification analysis of affine term structure models. Journal of Finance LV, 1943-1978.

16. Dai, Q. and Singleton, K.J. 2003. Term structure dynamics in theory and reality. Review of Financial Studies 16, 631-678.

17. Dai, Q., Singleton, K.J., Yang, W. 2007. Regime shifts in a dynamic term structure model of U.S. treasury bond yields. Review of Financial Studies 20(5), 1669-1706.

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

19. Dybvig, P.H., Ingersoll, J.E., Ross, S.A. 1996. Long forward and zero-coupon rates can never fall. Journal of Business 69, 1-25.

20. Esscher, F. 1932. On the probability function in the collective theory of risk. Skandinavisk Aktuarietidskrift 15, 175-195.

21. Elliott, R.J., Aggoun, L., Moore, J.B. 1994. Hidden Markov Models: estimation and control. Springer-Verlag, Berlin-Heidelberg-New York.

22. Elliott, R.J., J. van der Hoek. 2001. Stochastic flows and the forward measure. Finance and Stochastics 5, 511-525.

23. Elliott, R.J., Mamon R.S. 2003a. An interest rate model with a Markovian mean-reverting level. Quantitative Finance 2(6), 454-458.

24. Elliott, R.J., Mamon, R.S. 2003b. A complete yield curve descriptions of a Markov interest rate model. International Journal of Theoretical and Applied Finance 6(4), 317-326.

25. Elliott, R.J., P.E. Kopp. 2004. Mathematics of financial markets. Springer Verlag, Berlin-Heidelberg-New York.

26. Elliott, R.J., Chan, L., Siu, T.K. 2005. Option pricing and Esscher transform under regime switching. Annals of Finance 1(4), 423-432.

27. Elliott, R.J., Wilson, C.A. 2007. The term structure of interest rates in a hidden Markov setting. Hidden Markov Models in Finance (R.S. Mamon and R.J. Elliott, eds.). International Series in Operations Research & Management Science, Springer Verlag, Berlin-Heidelberg-New York, pp. 15-30.

28. Elliott, R.J., Siu, T.K. 2009. On Markov-modulated exponential-affine bond price formulae. Applied Mathematical Finance 16(1), 1-15.

29. Elliott, R.J., Siu, T.K. 2010. On risk minimizing portfolios under a Markovian regime-switching Black-Scholes economy. Annals of Operations Research 176(1), 271-291.

30. Esscher, F. 1932. On the probability function in the collective theory of risk. Skandinavisk Aktuarietidskrift 15, 175-195.

31. Evans, M.D.D. 1998. Regime shifts, risk and the term structure. Preprint. Department of Economics, Georgetown University.

32. Gerber, H.U., Shiu, E.S.W. 1994. Option pricing by Esscher transforms (with discussions). Transactions of the Society of Actuaries 46, 99-191.

33. Longstaff, F.A., Schwartz, E.S. 1992. Interest-rate volatility and the term structure: a two-factor general equilibrium model. The Journal of Finance 47, 1259-1282.

34. Merton, R.C. 1973. Theory of rational option pricing. Bell Journal of Economics and Management Science 4, 141-183.

35. Naik, V., Lee, M.H. 1997. Yield curves dynamics with discrete shifts in economic regimes: theory and estimation. Preprint. University of British Columbia.

36. Roma, A., Torous, W. 1997. The cyclical behavior of interest rates. Journal of Finance 52(4), 1519-1542.

37. Siu, T.K., Tong, H., Yang H. 2004. On pricing derivatives under GARCH models: a dynamic Gerber-Shiu's approach. North American Actuarial Journal 8(3), 17-31.

38. Siu, T.K. 2005. Fair valuation of participating policies with surrender options and regime switching. Insurance: Mathematics and Economics 37(3), 533-552.

39. Vasicek, O.A. 1977. An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177-188.

40. Yao, Y. 2002. State price density, Esscher transforms, and pricing options on stocks, bonds, and foreign exchange rates. North American Actuarial Journal 5(3), 104-117.