

2015-10-02

# Asset Pricing, Capital Structure and Financial Economics: Pricing, Hedging, and Risk

Shen, Jia

---

Shen, J. (2015). Asset Pricing, Capital Structure and Financial Economics: Pricing, Hedging, and Risk (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from

<https://prism.ucalgary.ca>. doi:10.11575/PRISM/25032

<http://hdl.handle.net/11023/2587>

*Downloaded from PRISM Repository, University of Calgary*

UNIVERSITY OF CALGARY

Asset Pricing, Capital Structure and Financial Economics: Pricing, Hedging, and Risk

by

Jia Shen

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN MANAGEMENT

CALGARY, ALBERTA

SEPTEMBER, 2015

© Jia Shen 2015

# Abstract

Asset Pricing is a central topic in Finance Theory. Explaining why different assets have different risk premia is a main goal of Asset Pricing Theory.

In this thesis, I explore various topics in Finance Theory, including topics in asset pricing, capital structure, and financial economics.

Chapter 1 investigates a regime switching Lucas economy in continuous time, with multiple dividend streams and labor income. We determine the asset prices in equilibrium in the economy with regime switching, and derive a system of partial differential equations for the asset prices and the short interest rate.

In Chapter 2, I consider credit risk. Motivated by empirical findings, we propose a framework using unobservable, underlying Markov chains, to model naturally both frailty and default contagion.

Chapter 3 is subsequent research based on Chapter 2. We consider a reduced-form, intensity-based credit risk model, which allows for both frailty and default contagion, using a so-called “self-exciting” intensity, in the sense that the default intensity varies not only with the risk factors, but also depends on the previous default history of all the firms.

In Chapter 4, we turn our attention to an area that is related to both asset pricing and corporate finance. We investigate the optimal capital structure of a corporate when the dynamics of the assets (both growth rate and volatility) change following different states of the economy. In Chapter 5, we investigate credit risk and the credit spread of a corporate defaultable bond when the dynamics of the assets change according to different states of the economy.

In Chapter 6, we investigate a model where the asset price follows *hidden* Markov modulated jump-diffusion dynamics. This framework incorporates two important empirically observed features: jumps and regime shifts.

# Acknowledgements

I would like to express the deepest and foremost gratitude to my supervisor Professor Robert Elliott. During my Ph.D., Robert is always there to help me. I virtually learned everything I know about stochastic calculus from him. It is my great fortune to have Robert as my advisor. He is always kind, nice, generous and patient to me, whenever I need help. Without his guidance, help, generosity and patience, this thesis would never have been possible.

I also thank the members in my supervisory committee. I learned a great deal of knowledge and idea from Dr. Alfred Lehar and Dr. K.J. Choi. I would express my appreciation to the Haskayne School of Business and the Finance Division for all the help and support that I obtained in the program.

I would like to thank my beloved wife, Jingjing Zhou, for her support and understanding. She has been always supportive and considerate to me. My family is one of the motivations that keep me moving forward.

# Table of Contents

<b>Abstract</b> . . . . .	i
<b>Acknowledgements</b> . . . . .	ii
Table of Contents . . . . .	iii
List of Tables . . . . .	vi
List of Figures . . . . .	vii
1 General Equilibrium Pricing with Multiple Dividend Streams and Regime Switching . . . . .	1
1.1 Introduction . . . . .	1
1.2 Model . . . . .	3
1.2.1 Dividends and Consumption . . . . .	3
1.2.2 Equity Market and Portfolio . . . . .	6
1.2.3 Preference . . . . .	7
1.2.4 General Equilibrium . . . . .	7
1.3 Solving the Optimization Problem . . . . .	8
1.3.1 Static Budget Constraint . . . . .	8
1.3.2 Solving Optimization Problem . . . . .	11
1.4 Asset Prices . . . . .	12
1.5 A System of Partial Differential Equations for the Asset Prices . . . . .	15
1.6 Risk Premium for Regime Switching . . . . .	16
1.6.1 Premium for Regime Switching Risk . . . . .	16
1.6.2 Equity Premium with Regime Switching . . . . .	22
1.7 Endogenous Interest Rate . . . . .	23
1.7.1 General Utility . . . . .	23
1.7.2 Bond Price and Yield . . . . .	24
1.7.3 Constant Relative Risk Aversion Utility . . . . .	26
1.8 Term Structure of Interest Rates . . . . .	27
1.9 Conclusion . . . . .	30
2 Credit Risk with Latent Contagion and Frailty: Default Probabilities, Pricing, and Hedging . . . . .	40
2.1 Introduction . . . . .	40
2.2 An Ideal Model Dynamics and Methodology . . . . .	44
2.2.1 An Ideal Model Dynamics with Contagion and Frailty . . . . .	44
2.2.2 Point Processes and Random Measure . . . . .	46
2.2.3 Dividend Streams . . . . .	47
2.2.4 Payoffs of Credit Derivatives . . . . .	50
2.3 Practical Models . . . . .	51
2.3.1 Model I: A Simple Model with One State Process . . . . .	51
2.3.2 Model II: A Practical Model with Contagion and Frailty . . . . .	54
2.3.3 Model III: A More Refined Model with Contagion and Frailty . . . . .	57
2.4 Default Probabilities . . . . .	59
2.5 Pricing of Contingent Claims . . . . .	62
2.5.1 Price of a Contingent Claim in Model I . . . . .	63

2.5.2	Price of a Contingent Claim in Model II . . . . .	65
2.5.3	Price of a Contingent Claim in Model III . . . . .	66
2.6	Dynamic Hedging . . . . .	67
2.6.1	Scheme of Hedging . . . . .	67
2.6.2	Hedging Strategies for the Credit Derivatives . . . . .	71
2.7	Parameter Estimation . . . . .	75
2.7.1	Parameter Estimation for Default Probabilities from Historical Defaults $Z$ . . . . .	75
2.7.2	Joint Parameter Estimation for the Pricing of Credit Securities . . . . .	77
2.8	Concluding Remarks . . . . .	79
3	Credit Risk and Contagion via Self-Exciting Default Intensity . . . . .	89
3.1	Introduction . . . . .	89
3.2	Model Dynamics and Methodology . . . . .	93
3.2.1	Model Dynamics with Contagion and Frailty . . . . .	94
3.2.2	Simulation of Default times . . . . .	96
3.2.3	Point Processes and Random Measure . . . . .	96
3.2.4	Dividend Streams . . . . .	98
3.2.5	Payoffs of Credit Derivatives . . . . .	99
3.3	Estimation of Current State . . . . .	100
3.3.1	Filter Equation . . . . .	100
3.3.2	A Further Filter Equation . . . . .	101
3.3.3	Robust Filters . . . . .	102
3.4	Pricing of Defaultable Contingent Claims . . . . .	104
3.4.1	Construction of General Dividends . . . . .	104
3.4.2	Price of a Defaultable Contingent Claim . . . . .	104
3.4.3	Examples: Some Credit Securities . . . . .	106
3.5	Parameter Estimation . . . . .	108
3.6	Concluding Remarks . . . . .	110
4	Dynamic Optimal Capital Structure with Regime Switching . . . . .	117
4.1	Introduction . . . . .	117
4.2	Asset Dynamics and Coupled PDE under Regime Switching . . . . .	120
4.2.1	Asset Dynamics under Regime Switching . . . . .	120
4.2.2	Coupled PDE for General Cash Flow Stream . . . . .	122
4.3	Simple Optimal Capital Structure with Regime Switching . . . . .	123
4.4	Bankruptcy Costs and Tax Effect . . . . .	125
4.4.1	Bankruptcy Costs . . . . .	126
4.4.2	Tax Benefits . . . . .	126
4.4.3	Value of the Firm . . . . .	126
4.4.4	A New Security for Smooth Pasting . . . . .	127
4.4.5	Optimal Capital Structure with Tax and Bankruptcy Costs under Regime Switching . . . . .	127
4.5	Concluding Remarks . . . . .	129
5	Credit Spread of Defaultable Corporate Bonds with Regime Switching . . . . .	142
5.1	Introduction . . . . .	142
5.2	Model with Default Only at Maturity . . . . .	145

5.3	Model with First-Passage-Time Default . . . . .	152
5.3.1	Coupled PDE for General Cash Flow Stream . . . . .	153
5.3.2	Defaultable Coupon Bond . . . . .	155
5.3.3	Perpetual Bond . . . . .	157
5.4	Concluding Remarks . . . . .	160
6	Asset Price and Option Pricing with Markov Modulated Jump-Diffusion Dynamics . . . . .	171
6.1	Introduction . . . . .	171
6.2	Model Specification . . . . .	172
6.3	Filtering of the Current State . . . . .	174
6.4	Parameter Estimation . . . . .	178
6.5	Applications on Option Pricing . . . . .	183
6.5.1	Idiosyncratic Jump Risk . . . . .	184
6.5.2	Systematic Jump Risk . . . . .	184
6.6	Concluding Remarks. . . . .	185
	<b>Appendix: List of Publications</b> . . . . .	187
	Bibliography . . . . .	188

## List of Tables

1.1	Premium for regime switching risk: parameters for the regime switching economy	19
1.2	Descriptive statistics for price-dividend ratios in economies . . . . .	20
1.3	Equity Premium: parameters for the regime switching economy . . . . .	23



# List of Figures and Illustrations

1.1	Premium for regime switching risk . . . . .	19
-----	---	----

# Chapter 1

## General Equilibrium Pricing with Multiple Dividend Streams and Regime Switching

### 1.1 Introduction

General equilibrium pricing in a consumption-based, endowment economy has been one of the central topics of asset pricing theory.

Many researchers have investigated the general equilibrium of an endowment economy. Lucas (1978) describes a discrete time, general equilibrium model with a representative agent where the economy's output is taken to be exogenously given, and derives the Euler equation for the price of the risky equity. Bakshi and Chen (1996) study a variant of the Lucas economy, assuming the representative agent obtains utility from both a consumption stream and money balance. Cecchetti, Lam and Mark (1993) modify the Lucas model in discrete time to match the equity premium. Heaton and Lucas (1996) consider a variation of Lucas model to allow transaction costs. Cochrane, Longstaff and Santa-Clara (2007) consider a Lucas economy with two dividend sources (trees) and infinite time horizon, using log utility. Similarly, Parlour, Stanton and Walden (2011) consider a Lucas economy with CRRA utility and two dividend sources, one of which is a risky dividend stream and the other is a risk-free dividend stream. Parlour, Stanton and Walden (2011) use this economy to examine the equity premium and volatility puzzles, and find that the model works quite well. Buraschi, Porchia and Trojani (2010) consider a Lucas model subject to persistent distress events. Martin (2013) examines the Lucas model with a CRRA utility and many dividend streams. All these papers use stochastic dynamic programming to find the optimal consumption and portfolio policies.

On the other hand, there has been a large body of empirical evidence suggesting that the aggregate economy is characterized by periodic shifts between distinct regimes of the business cycle.

In this paper, we investigate a regime switching Lucas economy in continuous time, with multiple dividend streams and allowing labor income. We consider a continuous time, pure exchange economy with a single consumption good and  $N$  dividend sources. We assume the economy has  $K$  states, representing “recession”, “expansion”, and other possible intermediate states. Thus, there are two types of shocks in this economy: individual shocks and systematic shocks. The individual shocks are modeled by the Brownian motions, which only affect the volatility of the individual dividend streams. The systematic shocks are modeled by a continuous time Markov chain  $Y$ , whose state shifts affect the expected return and the volatility of all the dividend streams. Thus, during an expansion state, the firms’ dividend streams grow more quickly and remain more stable, while during a recession state, the dividend streams have lower expected growth rate.

Our goal in this paper is threefold. Firstly, we model a regime switching Lucas economy in continuous time with multiple dividend streams. Most previous contributions concentrate on the case where there is only one dividend stream. Regime switching has not been investigated in a Lucas economy. Secondly, we examine the endogenous short interest rate and bond prices under regime switching. There is an equivalent martingale measure in this setting. Thirdly, we explore applications of our equilibrium model with regime switching. We show that, for example, the economy implies a rich framework for the term structure of interest rates. It can imply both affine and quadratic term structure.

Also, instead of the stochastic dynamic programming used in previous works, we apply the martingale/duality approach as developed in Karatzas, Lehocky and Shreve (1987), Cox and Huang (1989) and Pliska (1986).

The paper is organized as follows. In Sections 2 and 3, we describe our model, and

briefly review the concepts and methodology. In Section 4, we determine the asset prices in equilibrium in our economy with regime switching. In Section 5, we derive a system of partial differential equations for the asset prices and the short interest rate. In Section 6, we obtain the solution for the endogenous short interest rate, the bond price, and the yield of the bond. In Section 7, we consider the application of the equilibrium model and show that the model implies a rich framework for the term structure of interest rates.

## 1.2 Model

### 1.2.1 Dividends and Consumption

Consider an economy with a representative agent, which represents a large number of identical consumers. The economy has a single consumption good. The consumption good is produced by  $N$  productive units. The productions of the good are regarded as  $N$  exogenous dividend processes.

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, which is rich enough to imply the randomness of our economy. The instantaneous dividend stream is  $D_t = (D_t^1, D_t^2, \dots, D_t^N)$ . Assume the dynamics of  $D_t$  are governed by

$$dD(t) = \text{diag}(D_t) (\mu^D(t)dt + \sigma_t^D dW_t), \quad (1.1)$$

where  $W = \{W_t, 0 \leq t \leq T\}$  is an  $N$ -dimensional  $\mathbb{P}$ -Brownian motion. Or, in matrix form,

$$d \begin{bmatrix} D_t^1 \\ D_t^2 \\ \vdots \\ D_t^N \end{bmatrix} = \text{diag} \left( \begin{bmatrix} D_t^1 & D_t^2 & \dots & D_t^N \end{bmatrix} \right) \left( \begin{bmatrix} \mu_1^D(t) \\ \mu_2^D(t) \\ \vdots \\ \mu_N^D(t) \end{bmatrix} dt + \begin{bmatrix} \sigma_{11}^D(t) & \sigma_{12}^D(t) & \dots & \sigma_{1N}^D(t) \\ \sigma_{21}^D(t) & \sigma_{22}^D(t) & \dots & \sigma_{2N}^D(t) \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{N1}^D(t) & \sigma_{N2}^D(t) & \dots & \sigma_{NN}^D(t) \end{bmatrix} \begin{bmatrix} dW_t^1 \\ dW_t^2 \\ \vdots \\ dW_t^N \end{bmatrix} \right).$$

We suppose the state of the economy is described by a finite state continuous-time Markov chain  $Y = \{Y_t, t \geq 0\}$ . In our model, the states of the Markov Chain represent different economic environments. For example, there could be just two states for  $Y$  representing

“good” (expansion) and “bad” (recession) economic regimes. The switching of the states of the economy can be attributed to structural changes in macroeconomic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the productive units (dividend streams).

For notational convenience, the state space of  $Y$  can be taken to be, without loss of generality, the set of unit vectors  $\{e_1, e_2, \dots, e_N\}$ , where  $e_i = (0, \dots, 0, 1, 0, \dots, 0)' \in \mathbb{R}^N$  with the “1” in the  $i$ th entry and where the superscript  $'$  represents the transpose of a row vector. This is called the *canonical representation* of the state space of the Markov chain  $Y$ . Suppose the process  $Y$  is homogeneous in time. Let  $A$  denote the constant rate matrix<sup>1</sup> of the Markov chain  $Y$ . Then, using the canonical representation of the state space, the dynamics of the Markov chain  $Y$  have the following *semi-martingale representation* (see Elliott et al. (1995))

$$dY_t = AY_t dt + dM_t, \quad (1.2)$$

where  $(M_t)_{t \geq 0}$  is an  $\mathbb{R}^N$ -valued martingale with respect to the natural filtration generated by  $Y$ . From equation (1.2), an explicit solution for  $Y_t$  is

$$Y_t = \left( Y_0 + \int_0^t \exp(-As) dM_s \right) \exp(At).$$

In our regime switching economy, the drift and volatility coefficients in equation (1.1),  $\mu^D(t)$  and  $\sigma^D(t)$ , both depend on the state  $Y_t$  of the of the economy. Then, for  $1 \leq i \leq N$ ,

$$\mu_i^D(t) = \mu_i^D(Y_t).$$

We suppose there are constants  $\mu_{i,1}^D, \dots, \mu_{i,K}^D$  such that  $\mu_i^D(Y_t) = \mu_{i,k}^D$  if  $Y_t = e_k$ ,  $1 \leq k \leq K$ .

Write  $\mu_i^D = (\mu_{i,1}^D, \dots, \mu_{i,K}^D)$ . Then

$$\mu_i^D(t) = \mu_i^D(Y_t) = \langle \mu_i^D, Y_t \rangle = \sum_{k=1}^K \langle \mu_i^D, e_k \rangle \mathbf{1}_{\{Y_t=e_k\}}.$$

---

<sup>1</sup>Note that in our setting  $A$  is the transpose of the usual generator matrix of a Markov chain.

Similarly, we suppose there are constants  $\sigma_{ij,1}^D, \dots, \sigma_{ij,K}^D$  such that  $\sigma_{ij}^D(Y_t) = \sigma_{ij,k}^D$  if  $Y_t = e_k$ . Write  $\sigma_{ij}^D = (\sigma_{ij,1}^D, \dots, \sigma_{ij,K}^D)$ . Then

$$\sigma_{ij}^D(t) = \sigma_{ij}^D(Y_t) = \langle \sigma_{ij}^D, Y_t \rangle = \sum_{k=1}^K \langle \sigma_{ij}^D, e_k \rangle \mathbf{1}_{\{Y_t=e_k\}}.$$

The simplest example is when  $N = 1$ . Intuitively, if  $Y_0 = e_i$ , then the dividend process follow the dynamics

$$\frac{dD_t}{D_t} = \mu_{it}dt + \sigma_{it}dW_t.$$

The dividend  $D_t$  follows the above dynamics until the underlying Markov chain  $Y$  switches its state to some  $e_j$ , and this is when the dynamics of the dividend change to

$$\frac{dD_t}{D_t} = \mu_{jt}dt + \sigma_{jt}dW_t.$$

The dividend process continues to follow the above dynamics until the underlying Markov chain  $Y$  switches its state next time to some  $e_k$ , when new drift  $\mu_k$  and  $\sigma_k$  governs the dynamics, and so on. Therefore, a regime switching dividend process has its drift and volatility coefficients switch according to the state of the Markov chain.<sup>2</sup>

The representative agent's instantaneous consumption at time  $t$  is denoted  $c(t)$ . We suppose the representative agent has instantaneous earning  $e_t$  (e.g., labor income) at time  $t$ , and

$$de_t = e_t(\mu_t^e dt + \sigma_t^e dW_t^e),$$

where  $W^e$  is a Brownian motion independent of  $W$  (for dividend streams). The coefficients  $\mu_t^e$  and  $\sigma_t^e$  also switch following the Markov chain  $Y$ . Denote the agent's cumulative net surplus income stream by

$$I_t = \int_0^t (e_u - c_u) du.$$

---

<sup>2</sup>More detailed comparison of regime switching and non-switching economies from an agent's perspective is in Section 1.6.1.

## 1.2.2 Equity Market and Portfolio

The economy has a financial market, which has riskless asset (bond)  $S^0$ . The productive units (the dividend processes) issue equities in the financial market. For simplicity, we suppose each unit only issues one share of equity on the market. The price of the the dividend process  $D^i$  is  $S^i$ . Suppose

$$dS_t^i = S_t^i (\mu_i^S(t)dt + \sigma_i^S(t) \cdot dW_t), \quad i = 1, \dots, N,$$

where  $W$  is the  $N$ -dimensional Brownian motion for the dynamics of the dividend streams. In our equilibrium economy, the price  $S_t = (S_t^1, \dots, S_t^N)$  for the dividends is to be endogenously determined.

At each time  $t$ , investor holds a portfolio of the equities. Let  $N_t^i$  be the number of shares of equity  $i$  at time  $t$ , and  $\pi_t^i = N_t^i S_t^i$  denote the dollar amounts invested in asset  $i$  at time  $t$ ,  $i = 1, \dots, N$ .

In addition to the  $N$  risky assets, there is a riskless asset  $S^0$  in the market. The riskless asset  $S^0$  follows the dynamics

$$dS_t^0 = S_t^0 r_t dt,$$

where  $r_t$  is the riskless rate of return, or instantaneous (short) interest rate. We shall show that  $r_t$  is endogenously determined in equilibrium in our economy.

Write  $X_t$  for the agent's wealth level at time  $t$ . Then

$$X_t = \sum_{i=0}^N N_t^i S_t^i = \sum_{i=0}^N \pi_t^i.$$

The investor obtains gains from holding equities. The gains process  $\{G_t\}_{0 \leq t \leq T}$  from investments is given by

$$dG_t = \sum_{i=0}^N N_t^i (dS_t^i + D_t^i dt) = N_t^0 dS_t^0 + \sum_{i=1}^N N_t^i (dS_t^i + D_t^i dt),$$

and so

$$G_t = \sum_{i=0}^N \int_0^t N_u^i (dS_u^i + D_u^i du) = \int_0^t N_u^0 dS_u^0 + \sum_{i=1}^N \int_0^t N_u^i (dS_u^i + D_u^i du).$$

Note that  $X_t$  is also  $X_t = I_t + G_t$ .

### 1.2.3 Preference

The representative agent has preference for the consumption  $U(t, c_t) = e^{-\int_0^t \rho_s ds} u(c_t)$ , where  $\rho_s$  is the subjective discount factor at time  $s$ . In our Markov switching setting, the subjective discount factor  $\rho_t$  is allowed to vary with the states of the underlying Markov chain. Thus  $\rho_t = \rho(Y_t)$ . We assume there are constants  $\rho^{(1)}, \dots, \rho^{(K)}$  such that  $\rho(Y_t) = \langle \rho, Y_t \rangle$ .

First we derive the budget constraint for the investor. The dynamic budget constraint is

$$\begin{aligned} dX_t &= dI_t + dG_t \\ &= dI_t + r_t X_t dt + \pi \cdot \left\{ (\alpha^S(t) - r_t \cdot \mathbf{1}) dt + \sigma^S(t) dW_t \right\}, \end{aligned}$$

where

$$\begin{aligned} \alpha_i^S(t) &:= \mu_i^S(t) + \frac{D_t^i}{S_t^i}, \\ \pi &= \left[ \pi_t^1 \quad \pi_t^2 \quad \dots \quad \pi_t^N \right]. \end{aligned}$$

It can be shown that the solution of the budget constraint is (for a proof, see Appendix)

$$d \left( \frac{X_t}{S_t^0} \right) = \frac{\pi_t \cdot ((\alpha^S(t) - r_t \cdot \mathbf{1}) dt + \sigma^S(t) dW_t) + dI_t}{S_t^0} \quad (1.3)$$

or equivalently

$$\begin{aligned} \frac{X_t}{S_t^0} &= \frac{X_0}{S_0^0} + \int_0^t \frac{\pi_u}{S_u^0} \cdot \left( (\alpha_u - r_u \cdot \mathbf{1}) du + \sigma_u^S dW_u \right) + \int_0^t \frac{1}{S_u^0} dI_u \\ &= X_0 + \int_0^t \frac{\pi_u}{S_u^0} \cdot \left( (\alpha_u - r_u \cdot \mathbf{1}) du + \sigma_u^S dW_u \right) + \int_0^t \frac{1}{S_u^0} dI_u. \end{aligned}$$

The optimization problem of the representative agent is then

$$\begin{aligned} \sup_{(c, \pi)} \mathbb{E} \left[ \int_0^T e^{-\int_0^t \rho_s ds} u(c_t) dt \right] \\ \text{subject to } dX_t = dI_t + r_t X_t dt + \pi \left( (\alpha_t - r_t \cdot \mathbf{1}) dt + \sigma_t dW_t \right). \end{aligned} \quad (1.4)$$

### 1.2.4 General Equilibrium

The general equilibrium in this economy is defined as follows.



**Definition 1** *An equilibrium in the economy consists of processes  $\{c_t, \pi_t, S_t, r_t\}$  such that*

- *the representative agent solves the optimal problem*

$$\begin{aligned} & \sup_{(c, \pi)} \mathbb{E} \left[ \int_0^T e^{-\int_0^t \rho_s ds} u(c_t) dt \right] \\ & \text{s.t. } dX_t = dI_t + r_t X_t dt + \pi \left( (\alpha_t - r_t \cdot \mathbf{1}) dt + \sigma_t dW_t \right). \end{aligned} \tag{1.5}$$

- *the commodity market clears:  $c_t = e_t + \sum_{i=1}^N D_t^i$ ,  $0 \leq t \leq T$*
- *the stock market clears:  $\pi_t = S_t \iff N_t = 1$ , where for simplicity, we suppose the equity is given by only one share on the market.*

In this economy, the consumption and stock holdings are exogenously given. The interest rate  $r_t$  and the price of the dividend processes  $S_t$  are endogenous. We shall determine the interest rate  $r_t$  and the prices  $S_t$ .

### 1.3 Solving the Optimization Problem

In this section, we give the martingale/duality approach and solve the agent's optimization problem.

$$\begin{aligned} & \sup_{(c, \pi)} \mathbb{E} \left[ \int_0^T e^{-\int_0^t \rho_s ds} u(c_t) dt \right] \\ & \text{subject to } dX_t = dI_t + r_t X_t dt + \pi \left( (\alpha_t - r_t \cdot \mathbf{1}) dt + \sigma_t^S dW_t \right). \end{aligned} \tag{1.6}$$

#### 1.3.1 Static Budget Constraint

To solve the optimization problem of the representative agent, we first derive the static budget constraint. Since the solution of

$$dX_t = dI_t + r_t X_t dt + \pi \left( (\alpha_t - r_t \cdot \mathbf{1}) dt + \sigma_t^S dW_t \right)$$

is given by (as shown in Appendix)

$$d \left( \frac{X_t}{S_t^0} \right) = \frac{\pi_t}{S_t^0} \cdot \left( (\alpha_t - r_t \cdot \mathbf{1}) dt + \sigma_t^S dW_t \right) + \frac{1}{S_t^0} dI_t,$$

we consider a new measure  $\tilde{\mathbb{P}}$ . Define  $\theta_t = (\theta_t^1, \dots, \theta_t^N)^\top$  by

$$\alpha_i^S(t) - r_t = \sum_{j=1}^N \sigma_{ij}^S(t) \theta_t^j, \quad i = 1, \dots, m,$$

or equivalently,

$$\theta = (\sigma^S)^{-1} (\alpha^S(t) - r_t \cdot \mathbf{1}).$$

Consider

$$Z_t = \exp \left( - \int_0^t \theta_u \cdot dW_u - \frac{1}{2} \int_0^t \|\theta_u\|^2 du \right).$$

Then Itô's Lemma gives

$$dZ_t = -Z_t(\theta_t \cdot dW_t).$$

Thus  $\mathbb{E}[Z_T] = 1$ . Define the new probability measure  $\tilde{\mathbb{P}}$  by

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) \quad \text{for all } A \in \mathcal{F}.$$

Define  $\tilde{W}_t$  by

$$d\tilde{W}_t = \theta_t + dW_t.$$

The Girsanov Theorem implies that under the new measure  $\tilde{\mathbb{P}}$ , the process  $\tilde{W}$  is a Brownian motion. Then under the new measure  $\tilde{\mathbb{P}}$ ,

$$\begin{aligned} \frac{X_t}{S_t^0} - X_0 - \int_0^t \frac{1}{S_u^0} dI_u &= \int_0^t \frac{\pi_u}{S_u^0} \cdot ((\alpha^S(u) - r_u \cdot \mathbf{1}) du + \sigma_u dW_u) \\ &= \int_0^t \frac{\pi_u^1}{S_u^0} \sigma_u d\tilde{W}_u \end{aligned}$$

is a local martingale. Define a density process

$$\zeta_t = \frac{Z_t}{S_t^0}.$$

Consider  $\zeta_t X_t$ . Using product rule, we have

$$\begin{aligned} d(\zeta_t X_t) &= d \left( \frac{Z_t}{S_t^0} X_t \right) = d \left( Z_t \frac{X_t}{S_t^0} \right) \\ &= Z_t d \left( \frac{X_t}{S_t^0} \right) + d(Z_t) \frac{X_t}{S_t^0} + (dZ_t) d \left( \frac{X_t}{S_t^0} \right) \end{aligned}$$

$$\begin{aligned}
&= Z_t \left( \frac{\pi_t}{S_t^0} \cdot ((\alpha_t - r_t \cdot \mathbf{1})dt + \sigma_t^S dW_t) + \frac{1}{S_t^0} dI_t \right) - Z_t(\theta_t \cdot dW_t) \frac{X_t}{S_t^0} \\
&\quad - Z_t(\theta_t \cdot dW_t) \left( \frac{\pi_t}{S_t^0} \cdot ((\alpha_t - r_t \cdot \mathbf{1})dt + \sigma_t^S dW_t) + \frac{1}{S_t^0} dI_t \right) \\
&= \zeta_t \pi_t \sigma_t^S (\theta_t dt + dW_t) + \zeta_t dI_t - \zeta_t X_t (\theta_t \cdot dW_t) - \zeta_t (\theta_t \cdot dW_t) \pi_t \sigma_t^S (\theta_t dt + dW_t) + \zeta_t (\theta_t dW_t) dI_t \\
&= \zeta_t \pi_t \sigma_t^S (\theta_t dt + dW_t) + \zeta_t dI_t - \zeta_t X_t (\theta_t dW_t) - \zeta_t (\pi_t \sigma_t) \theta_t dt \\
&= \zeta_t \pi_t \sigma_t^S \theta_t dt + \zeta_t \pi_t \sigma_t dW_t + \zeta_t dI_t - \zeta_t X_t (\theta_t dW_t) - \zeta_t \pi_t \sigma_t \theta_t dt \\
&= \zeta_t dI_t + \zeta_t (\pi_t \sigma_t^S - X_t \theta_t) dW_t.
\end{aligned}$$

Thus

$$\begin{aligned}
\zeta_t X_t &= \zeta_0 X_0 + \int_0^t \zeta_u dI_u + \int_0^t \zeta_u (\pi_u \sigma_u - X_u \theta_u) dW_u \\
&= X_0 + \int_0^t \zeta_u dI_u + \int_0^t \zeta_u (\pi_u \sigma_u - X_u \theta_u) dW_u, \quad 0 \leq t \leq T.
\end{aligned}$$

We see that

$$\zeta_t X_t - X_0 - \int_0^t \zeta_u dI_u = \int_0^t \zeta_u (\pi_u \sigma_u - X_u \theta_u) dW_u$$

is a local martingale under  $\mathbb{P}$ .

Since  $dI_t = (e_t - c_t)dt$ , we have

$$\zeta_t X_t - \int_0^t \zeta_u (e_u - c_u) du = X_0 + \int_0^t \zeta_u (\pi_u \sigma_u - X_u \theta_u) dW_u,$$

for which the right-hand side is a local martingale. The left-hand side of the above equation is bounded from below, thus it is a supermartingale. Then

$$\mathbb{E} \left[ \zeta_t X_t - \int_0^t \zeta_u (e_u - c_u) du \right] = X_0 + \mathbb{E} \left[ \int_0^t \zeta_u (\pi_u \sigma_u - X_u \theta_u) dW_u \right] \leq X_0 + 0.$$

Thus the static budget constraint is

$$\mathbb{E} \left[ \zeta_T X_T - \int_0^T \zeta_u (e_u - c_u) du \right] \leq X_0 = x.$$

The optimization problem is, therefore, transformed to the problem

$$\begin{aligned}
&\sup_{(c, \pi)} \mathbb{E} \left[ \int_0^T e^{-\int_0^t \rho_s ds} u(c_t) dt \right] \\
&\text{subject to } \mathbb{E} \left[ \zeta_T X_T - \int_0^T \zeta_t (e_t - c_t) dt \right] \leq X_0 = x.
\end{aligned} \tag{1.7}$$

### 1.3.2 Solving Optimization Problem

We solve the optimization problem with the static budget constraint. Here we consider a more general optimization problem, where the investor obtains utility from both consumption stream  $u(c_t)$  and terminal wealth level  $v(X_T)$ , i.e.,

$$\begin{aligned} & \sup_{(c, \pi)} \mathbb{E} \left[ \int_0^T e^{-\int_0^t \rho_s ds} u(c_t) dt + e^{-\int_0^T \rho_s ds} v(X_T) \right] \\ & \text{subject to } \mathbb{E} \left[ \zeta_T X_T - \int_0^T \zeta_t (e_t - c_t) dt \right] \leq X_0 = x. \end{aligned} \quad (1.8)$$

Clearly, this general version contains the case where the investor obtains utility only from the consumption stream  $u(c_t)$ . Set up the Lagrangian

$$\mathcal{L} = \mathbb{E} \left[ \int_0^T e^{-\int_0^t \rho_s ds} u(c_t) dt + e^{-\int_0^T \rho_s ds} v(X_T) \right] + \lambda \left( x - \mathbb{E} \left[ \zeta_T X_T - \int_0^T \zeta_t (e_t - c_t) dt \right] \right).$$

Then the first order conditions give  $\partial \mathcal{L} / \partial c_t = 0$ ,  $\partial \mathcal{L} / \partial X_T = 0$  and  $\partial \mathcal{L} / \partial \lambda = 0$ , which imply

$$\begin{aligned} e^{-\int_0^t \rho_s ds} u'(c_t) &= \lambda \zeta_t, \\ e^{-\int_0^T \rho_s ds} v'(X_T) &= \lambda \zeta_T, \\ \mathbb{E} \left[ \zeta_T X_T - \int_0^T \zeta_t (e_t - c_t) dt \right] &= x. \end{aligned}$$

Denote by  $I_1$  and  $I_2$  the inverse functions of  $u'(c_t)$  and  $v'(X_T)$ , respectively. We get from the first two equations the solutions  $c_t$  and  $X_T$  in terms of  $\lambda$ :

$$\begin{aligned} c_t &= I_1(\lambda \zeta_t), \\ X_T &= I_2(\lambda \zeta_T). \end{aligned}$$

Substituting these expressions  $c_t$  and  $X_T$  into the third equations, we have

$$\mathbb{E} \left[ \zeta_T I_2(t, \lambda \zeta_T) + \int_0^T \zeta_t I_1(t, \lambda \zeta_t) dt \right] = x.$$

Write  $\mathcal{X}(\lambda) := \mathbb{E} \left[ \zeta_T I_2(t, \lambda \zeta_T) + \int_0^T \zeta_t I_1(t, \lambda \zeta_t) dt \right]$ . Then the above equation has a unique solution

$$\lambda^* = \mathcal{Y}(x) = \mathcal{X}^{-1}(x).$$

Substituting  $\lambda^*$  into  $c_t$  and  $X_T$ , we obtain

$$\begin{aligned} c_t &= I_1(\lambda^* \zeta_t), \\ X_T &= I_2(\lambda^* \zeta_T) \\ \lambda^* &= \mathcal{Y}(x). \end{aligned}$$

**Remark.** When  $v(\cdot) \equiv 0$ , so the agent obtains utility only from consumption, not from the terminal wealth, we see that  $X_T = 0$ .

## 1.4 Asset Prices

The standard Euler equation for the pricing formula is as follows.

### Proposition 1.4.1

$$e^{-\int_0^t \rho_s ds} u'(c_t) S_t = \mathbb{E}_t \left[ \int_t^T e^{-\int_0^s \rho_u du} u'(c_s) D_s ds \right]. \quad (1.9)$$

*Proof.* The proof is the same as in Duffie and Epstein (1992). Also, Duffie and Zame (1989), Duffie and Skiadas (1994), and Davis (1998) provide proofs of the Euler equation in slightly different approach.  $\square$

Then, for the general utility  $u(\cdot)$ , we can derive the stochastic discount factor (also referred to as state price density, state price deflator, or pricing kernel) as follows.

### Proposition 1.4.2

$$S_t^i = \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \frac{u'(e_s + \sum_{i=1}^N D_s^i)}{u'(e_t + \sum_{i=1}^N D_t^i)} D_s^i ds \right]$$

*Proof.* From equation (1.9) we have, for any fixed  $t > 0$ ,

$$\begin{aligned} S_t^i &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \frac{u'(c_s)}{u'(c_t)} D_s^i ds \right] \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \frac{u'(e_s + \sum_{i=1}^N D_s^i)}{u'(e_t + \sum_{i=1}^N D_t^i)} D_s^i ds \right]. \end{aligned}$$

Here the last equality is by the market clearing condition in equilibrium.  $\square$

**Remark.** Proposition 1.4.2 gives the stochastic discount factor,  $e^{-\int_t^s \rho_u du} \frac{u'(e_s + \sum_{i=1}^N D_s^i)}{u'(e_t + \sum_{i=1}^N D_t^i)}$ .

In the following proposition, we consider the *constant relative risk aversion* (CRRA) utility  $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ .

**Proposition 1.4.3** *Suppose the agent has CRRA utility, i.e.,  $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ . Then*

$$S_t^i = D_t^i \mathbb{E}_t \left[ \int_t^T \left( \frac{e_s + \sum_{i=1}^N D_s^i}{e_t + \sum_{i=1}^N D_t^i} \right)^{-\gamma} \exp \left( - \int_t^s \rho_u du \right) \exp \left( \int_t^s \mu_i^D(u) - \frac{1}{2} \|\sigma_i^D(u)\|^2 du + \int_t^s \sigma_i^D(u) \cdot dW_u \right) \right]$$

*Proof.* See Appendix.  $\square$

In general, one cannot derive an explicit analytic solution for  $S_t^i$  from the above expression. For  $N = 1$ ,  $e_t \equiv 0$ , and CRRA utility, we have an explicit expression for  $S_t$ .

**Theorem 1.4.4** *If  $N = 1$ ,  $e_t \equiv 0$ , and the agent has CRRA utility, i.e.,  $u(C_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ , then*

$$S_t = D_t \int_t^T \langle e^{(A + \text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds, \quad (1.10)$$

Here  $(\text{diag } \lambda) = \text{diag}(\lambda_1, \dots, \lambda_K)$ , where

$$\lambda_i = -\rho_i + (1-\gamma)\mu_i^D - \frac{1}{2}\gamma(1-\gamma)(\sigma_i^D)^2 = \left\langle -\rho + (1-\gamma)\mu^D - \frac{1}{2}\gamma(1-\gamma)(\sigma^D)^2, e_i \right\rangle.$$

*Proof* See Appendix  $\square$

Note that, if there is no regime switching, i.e., the state process  $Y$  has only one state, then  $\sigma_u^D$ ,  $\mu_u^D$ ,  $\lambda_u$ , are constants. In this case, we have a particularly simple expression for  $S_t$ .

**Corollary 1.4.5** *If there is no regime switching, i.e., the state process  $Y$  has only one state, then  $\sigma_u^D$ ,  $\mu_u^D$ ,  $\lambda_u$ , are constants, and*

$$S_t = \frac{D_t}{\lambda} (e^{\lambda(T-t)} - 1), \quad (1.11)$$

where  $\lambda = -\rho + (1-\gamma)\mu^D - \frac{1}{2}\gamma(1-\gamma)(\sigma^D)^2$ .

*Proof.* We know from Theorem 1.4.4 that

$$\begin{aligned}
S_t &= D_t \mathbb{E} \left[ \int_t^T \exp \left( \int_t^s \lambda_u du \right) ds \middle| \mathcal{F}_t \right] \\
&= D_t \mathbb{E} \left[ \int_t^T \exp (\lambda(s-t)) ds \middle| \mathcal{F}_t \right] \\
&= \frac{D_t}{\lambda} (e^{\lambda(T-t)} - 1).
\end{aligned}$$

□

**Remark.** Suppose starting at time  $t$  in Theorem 1.4.4, the regime switching effect vanishes, i.e., the economy becomes a non-switching economy from time  $t$  onwards. Then  $A = 0$ , and the price  $S_t$  is, provided the state of the economy at time  $t$  is  $Y_t = e_i$ ,

$$\begin{aligned}
S_t &= D_t \int_t^T \langle e^{(\text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds \\
&= D_t \int_t^T \langle e^{(\text{diag } \lambda)(s-t)} e_i, \mathbf{1} \rangle ds.
\end{aligned} \tag{1.12}$$

It is easy to verify that equation (1.12) can be simplified to

$$S_t = \frac{D_t}{\lambda_i} (e^{\lambda_i(T-t)} - 1). \tag{1.13}$$

This is consistent with equation (1.11) in Corollary 1.4.5, where the price  $S_t$  is derived directly.<sup>3</sup>

When the time horizon is infinite, i.e.,  $T \rightarrow \infty$ ,  $S_t$  is given by the following simpler formula.

**Corollary 1.4.6** *Suppose there is no regime switching, i.e., the state process  $Y$  has only one state. When  $T \rightarrow \infty$ , if  $\lambda < 0$ , we have*

$$S_t = \frac{D_t}{\rho - (1 - \gamma)\mu^D + \frac{1}{2}\gamma(1 - \gamma)(\sigma^D)^2}.$$

---

<sup>3</sup>Equations (1.12) and (1.13) are useful when we estimate the premium for the regime switching risk in Section 1.6.1.

*Proof.*

$$\begin{aligned}
S_t &= \frac{D_t}{\lambda} e^{\lambda u} \Big|_{u=0}^{T-t} \\
&= -\frac{D_t}{\lambda} \\
&= \frac{D_t}{\rho - (1 - \gamma)\mu^D + \frac{1}{2}\gamma(1 - \gamma)(\sigma^D)^2}.
\end{aligned}$$

□

**Remark.** When there is no regime switching, i.e., the state process has only one state, an equivalent approach to calculate  $S_t$  is as follows. This expression is also obtained in Corollary 1.5.2.

$$\begin{aligned}
S_t &= \frac{D_t}{\mu - \theta\sigma - r} (e^{(\mu - \theta\sigma - r)(T-t)} - 1) \\
&= \frac{D_t}{\mu - (\gamma\sigma_t^D)\sigma_t - \rho - \gamma\mu_t^D + \frac{1}{2}\gamma(\gamma + 1)(\sigma_t^D)^2} \\
&\quad \times \left( \exp \left( (\mu - (\gamma\sigma_t^D)\sigma_t - \rho - \gamma\mu_t^D + \frac{1}{2}\gamma(\gamma + 1)(\sigma_t^D)^2)(T - t) \right) - 1 \right).
\end{aligned}$$

## 1.5 A System of Partial Differential Equations for the Asset Prices

The dynamic budget constraint is

$$dX_t = dI_t + r_t X_t dt + \pi \left( (\alpha_t - r_t \cdot \mathbf{1}) dt + \sigma_t^S dW_t \right)$$

Suppose there exists a function  $f_i(x)$  such that

$$S_t^i = f_i(D_t).$$

**Theorem 1.5.1** *For  $i = 1, 2, \dots, N$ ,  $S_t^i = f_i(D_t)$  satisfies the following system of partial differential equations*

$$r_t f_i(D_t^i) = D_t^i + D_t^i \mu_i^D(t) f_i'(D_t^i) - \left( \sum_{j=1}^N \sigma_{ij}^S(t) \theta_t^j \right) f_i(D_t^i) + \frac{1}{2} f_i''(D_t^i) (D_t^i)^2 \|\sigma_t^D\|^2.$$



*Proof.* See Appendix. □

Solving this system of partial differential equations, we can obtain the equilibrium prices  $S_t^1, \dots, S_t^N$ . In particular, when there is only one dividend stream, the analytic solution has the following simple form.

**Corollary 1.5.2** *When  $N = 1$ , we have*

$$\begin{aligned} f(D_t) &= \frac{D_t}{\mu^D - \theta\sigma^D - r} \left( e^{(\mu^D - \theta\sigma^D - r)(T-t)} - 1 \right) \\ &= \frac{D_t}{\rho - (1 - \gamma)\mu^D + \frac{1}{2}\gamma(1 - \gamma)(\sigma^D)^2} \\ &\quad \times \left( \exp \left( (\rho - (1 - \gamma)\mu^D + \frac{1}{2}\gamma(1 - \gamma)(\sigma^D)^2)(T - t) \right) - 1 \right). \end{aligned}$$

*Proof.* When  $N = 1$ , we have the single partial differential equation

$$\theta_t \sigma_t^S = \frac{f'(D_t) \mu_t^D D_t + \frac{1}{2} f''(D_t) (\sigma_t^D)^2 D_t^2}{f(D_t)} - r_t + \frac{D_t}{f(D_t)},$$

or equivalently,

$$r_t f(D_t) = D_t + \mu_t^D D_t f'(D_t) - \theta_t \sigma_t^D f(D_t) + \frac{1}{2} f''(D_t) (\sigma_t^D)^2 D_t^2.$$

Solving this partial differential equation for  $f$ , we obtain

$$\begin{aligned} f(D_t) &= \frac{D_t}{\mu^D - \theta\sigma^D - r} \left( e^{(\mu^D - \theta\sigma^D - r)(T-t)} - 1 \right) \\ &= \frac{D_t}{\rho - (1 - \gamma)\mu^D + \frac{1}{2}\gamma(1 - \gamma)(\sigma^D)^2} \\ &\quad \times \left( \exp \left( (\rho - (1 - \gamma)\mu^D + \frac{1}{2}\gamma(1 - \gamma)(\sigma^D)^2)(T - t) \right) - 1 \right). \end{aligned}$$

□

## 1.6 Risk Premium for Regime Switching

### 1.6.1 Premium for Regime Switching Risk

An immediate question is whether the regime switching risk is priced. In this section we examine the risk premium of the regime switching risk.

First of all, since the representative agent's preference is modeled as an additively time-separable expected utility, the agent is risk averse (in mathematical terminology, the utility  $u(\cdot)$  is concave).<sup>4</sup> Since the agent is risk averse, the regime switching risk is naturally priced, as with extra uncertainty (regime shifts), the agent obtains less utility than the non-switching consumption.

From Theorem 1.4.4 we have

$$S_t = D_t \int_t^T \langle e^{(A+\text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds.$$

When  $A = 0$ , we obtain the asset price in the non-switching economy, as is shown in Corollary 1.4.5

$$\begin{aligned} S_t &= D_t \int_t^T \langle e^{(\text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds \\ &= D_t \text{diag} \left( \frac{1}{\lambda_1} (e^{\lambda_1(T-t)} - 1), \dots, \frac{1}{\lambda_1} (e^{\lambda_1(T-t)} - 1) \right) Y_t. \end{aligned}$$

The risk premium for regime switching is

$$D_t \int_t^T \langle (e^{(A+\text{diag } \lambda)(s-t)} - e^{(\text{diag } \lambda)(s-t)}) Y_t, \mathbf{1} \rangle ds.$$

More explicitly, we use linear approximation to obtain<sup>5</sup>

$$\begin{aligned} S_t &= D_t \int_t^T \langle e^{(A+\text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds \\ &\approx D_t \int_t^T \langle ((I + A + \text{diag } \lambda)(s-t)) Y_t, \mathbf{1} \rangle ds \\ &= D_t \int_t^T \langle ((I + \text{diag } \lambda)(s-t)) Y_t, \mathbf{1} \rangle ds + D_t \int_t^T \langle (A(s-t)) Y_t, \mathbf{1} \rangle ds \\ &\approx D_t \int_t^T \langle e^{(\text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds + D_t \int_t^T \langle (A(s-t)) Y_t, \mathbf{1} \rangle ds. \end{aligned} \tag{1.14}$$

---

<sup>4</sup>The utility function for consumption in general economic theory is explored in Mas-Colell et al. (1995). More recent textbooks that give excellent treatments for utility theory in asset pricing include LeRoy and Werner (2001), Pennacchi (2007), Lengwiler (2004), Altug and Labadie (2008). Back (2010) and Skiadas (2009) describe utility in the multiperiod discrete time and continuous time settings.

<sup>5</sup>Since the closed form solution for the regimes switching economy in Theorem 1.4.4 involves an integral for which the integrand is a matrix exponential, we are not able to obtain a simple, exact expression for the premium of the regime switching risk. Instead we resort to linear approximation to obtain estimated risk premium for regime switching effect.

In the above approximate expression, we can see that the first term is the price of the asset in the single-regime economy, and the second term can be interpreted as the approximate premium for the regime switching risk. Therefore, using the above crude linear approximation, the risk premium for the regime switching effect is

$$D_t \int_t^T \langle (e^{(A+\text{diag } \lambda)(s-t)} - e^{(\text{diag } \lambda)(s-t)}) Y_t, \mathbf{1} \rangle ds \approx D_t \int_t^T \langle (A(s-t)) Y_t, \mathbf{1} \rangle ds.$$

Using higher order approximations, we can obtain more accurate approximate solutions for the risk premium of the regime switching uncertainty.<sup>6</sup> The premium for the regime switching risk is determined by the rate matrix  $A$  of the regime switching process  $Y$ , and  $\text{diag}(\lambda)$  which is by definition determined by the switching drift  $\mu = (\mu_1, \dots, \mu_K)$ , and the switching volatility  $\sigma = (\sigma_1, \dots, \sigma_K)$ .<sup>7</sup>

To illustrate this more clearly, we conduct the following numerical analysis. Suppose the regime switching Lucas economy has two states,  $\{G, B\}$ , representing expansion and recession periods, respectively. Accordingly, the underlying Markov chain  $Y$  has two states,  $e_1 = (1, 0)'$  represents state  $G$  (expansion), and  $e_2 = (0, 1)'$  represents state  $B$  (recession), in the Lucas economy. The rate matrix of the underlying Markov chain is assumed to be

$$A = \begin{bmatrix} -0.3 & 0.5 \\ 0.3 & -0.5 \end{bmatrix}.$$

Extensive empirical studies document that (see, e.g., Mehra and Prescott (1985, 2003)) for the US economy since 1890's, the mean growth rate of consumption is around 1.018, and the standard deviation of the growth rate of consumption is 0.036. We follow Mehra and

---

<sup>6</sup>For example, we can use second order approximation to obtain a better estimate of the risk premium for regime switching. The expression involves both  $A$  and  $\lambda$ .

<sup>7</sup>In Dai, Singleton and Yang (2007), regime switching risk is priced through a parameterized specification of the pricing kernel, as a utility function is not used in their term structure model. The pricing measure is obtained by specifying the pricing kernel, which includes an additional component to explicitly account for the regime switching risk. In our general equilibrium framework, the agent's preference to risk is represented by the utility function, and the agent's risk aversion is characterized by the concavity of the utility function. Thus, the price of the regime switching risk is already captured by the pricing kernel, which is the intertemporal marginal rate of substitution (MRS) of consumption. By using the MRS as our pricing kernel, the regime switching risk is naturally priced. This is determined by agent's risk aversion through the concave, time-additive expected utility.

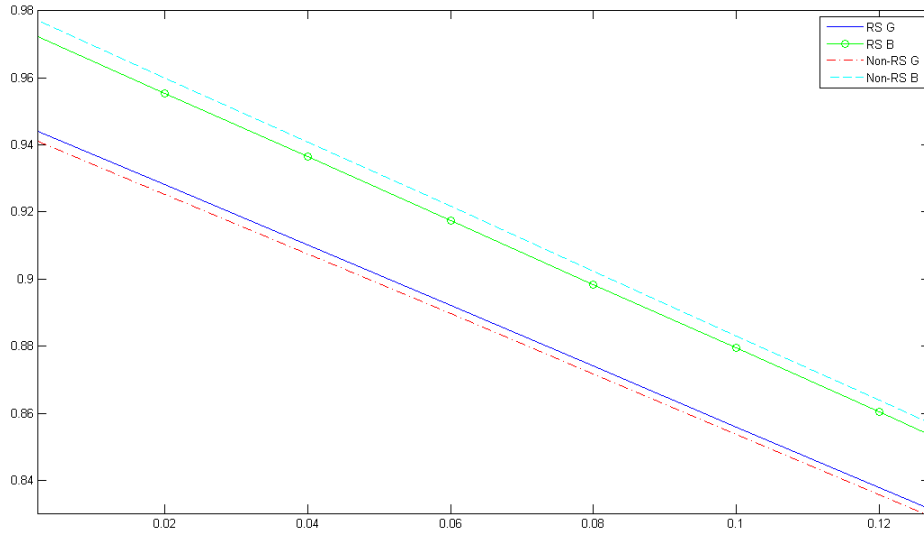


Figure 1.1: Premium for regime switching risk

Prescott (1985, 2003) to set parameters in our framework. We set  $(\rho_1, \rho_2) = (-\log 0.98, -\log 0.99) = (0.0202, 0.0101)$ , so that the agent's subjective discount factor is  $(e^{-\rho_1}, e^{-\rho_2}) = (0.98, 0.99)$  according to whether the current economy state is expansion or recession. A greater subjective discount factor  $e^{-\rho}$  means that the agent is more conservative on consumption, which is the case in a recession state. We set the risk aversion coefficient  $\gamma = 3$ . Set the growth rate of consumption  $(\mu_1, \mu_2) = (.05, .018)$ , the standard deviation of the consumption growth to be  $(\sigma_1^D, \sigma_2^D) = (.025, .036)$ . Thus, during an expansion state, the dividend streams grow more quickly and remain more stable, while during a recession state, the dividend streams have lower expected growth rate and are more volatile. The parameters are listed in Table 1.1.

parameter	state G (expansion)	state B (recession)
$\gamma$	3	3
$\rho$	0.0202	0.0101
$e^{-\rho}$	0.98	0.99
$\mu$	0.05	0.018
$\sigma_c$	0.025	0.036

Table 1.1: Premium for regime switching risk: parameters for the regime switching economy

We calculate the price-dividend ratios for regime-switching and non-switching economies, respectively, using equation (1.10) in Theorem 1.4.4, and equations (1.11) or (1.13) in Corollary 1.4.5 and its remark. In our calculation, the time horizon is taken to be  $T = 1$  (one year). We obtain four time series for the price-dividend ratio, for  $t \in [0, T]$ , for the regime-switching and non-switching economies with current state being expansion and recession, respectively. These four time series of price-dividend ratios are plotted in Figure 1.1, and their summary statistics are listed in Table 1.2.

In Figure 1.1, the lower-left curve (red curve) indicates the price-dividend ratio at time  $t$ , for the non-switching model when the economy is in an expansion state; the upper-right curve (cyan curve) indicates the price-dividend ratio at time  $t$ , for the non-switching model when the economy is in a recession state; the curve next to the lower-left (blue curve) indicates the price-dividend ratio at time  $t$ , for the regime switching model when the economy is in an expansion state; the curve next to the upper-right curve (green curve) indicates the price-dividend ratio at time  $t$ , for the regime switching model when the economy is in a recession state.

economy state	min	max	mean	std	skewness	kurtosis
Regime switching G (expansion)	0.019976	0.9460	0.4815	0.2812	-0.0385	1.8031
Regime switching B (recession)	0.0200	0.9742	0.4916	0.2898	-0.0202	1.7986
Non-switching G (expansion)	0.019976	0.9431	0.4807	0.2804	-0.0417	1.8010
Non-switching B (recession)	0.0200	0.9792	0.4930	0.2911	-0.0149	1.7993

Table 1.2: Descriptive statistics for price-dividend ratios in economies

Now we explain the intuition as to how the four types of economies are ranked by the agent's risk-averse preference. Intuitively, from an agent's perspective, a non-switching expansion state economy would be most favorable. Such a non-switching expansion economy typically has a high expected rate of return and low volatility, and the economy remains in the expansion state forever. Thus the agent will be best-off in such an economy. The worst economy would be a non-switching, recession state economy, which typically has a low expected rate of return and high volatility, and the economy will not change to a better state.

The next to best economy would be a regime switching economy which is in an expansion state. Such an economy typically has high expected rate of return and low volatility as it current stands, but bears a regime switching risk of changing to a recession state, where the economy turns to have low expected rate of return and high volatility, until the economy changes to an expansion state over again. Worse than a regime switching expansion state economy, but better than the recession state non-switching economy, is a regime switching recession state economy. Such an economy has a low expected rate of return and high volatility at its present state (recession), but will change to an expansion state, where the economy has a high expected rate of return and low volatility. From an agent's perspective, this is apparently better than remaining in a recession state forever.

This intuition explains the numerical analysis of the price-dividend ratio plotted in Figure 1.1. We can imagine that, at time  $t$ ,

- 1) if the dividend (consumption) stream bears no regime switching risk, and the economy is in an expansion state, then the expected rate of return, or asset price, that the agent requires will be the lowest (as represented by the red curve at lower-left in Figure 1.1);
- 2) if the dividend (consumption) stream bears the regime switching risk, and the economy is in an expansion state, then the expected rate of return, or the asset price, that the agent requires will be higher than that of a dividend stream without the regime switching risk, as the agent requires a premium for the additional regime switching risk (this is represented by the blue curve in Figure 1.1);
- 3) if the dividend (consumption) stream bears the regime switching risk, and the economy is in a recession state, then the expected rate of return, or the asset price, that the agent requires (green curve in Figure 1.1) will be higher than that of a regime switching, expansion state dividend stream (blue curve in

Figure 1.1), as the agent needs a premium for the low expected growth rate  $\mu^D$  and high volatility  $\sigma^D$ ; on the other hand, the expected rate of return, or the asset price, that the agent requires will be lower than that of a non-switching, recession state dividend stream (the cyan curve at upper-right in Figure 1.1), as the regime switching dividend stream will turn to an expansion state some time later (timing is determined by the process  $Y$ ), while the non-switching one stays in the recession state forever.

- 4) if the dividend (consumption) stream bears no regime switching risk, and the economy is in a recession state, then the expected rate of return, or the asset price, that the agent requires would be highest, as there is no hope that the dividend (consumption) stream will become better, and thus the agent requires the maximum expected rate of return.

## 1.6.2 Equity Premium with Regime Switching

In this section, we demonstrate the flexibility of the Lucas economy with regime switching. The empirical research for the US economy shows that the average annual real excess rate of return on the stock index (e.g., S&P 500) is about 8%. From the non-switching pricing formula, the excess rate of return for the risky asset is

$$\mu_t + \delta_t - r_t = \gamma \sigma_t^D \sigma_t^S \text{corr}(dc/c, dS/S)$$

Previous research shows that the model-implied expected excess rate of return on the market portfolio should be  $5 \times 0.2 \times 0.02 \times 0.2 = 0.4\%$ . This is the famous *equity premium puzzle*, which is one of the central topics in the modern asset pricing theory.

We shall demonstrate that, in our regime switching Lucas economy, the model-implied expected excess rate of return on the market can be improved in order of magnitudes.

We adapt the parameters used by previous researchers. Suppose for  $t = 1$ , the one-year

Parameter	State G (expansion)	State B (recession)
$\gamma$	5	5
$\sigma_t^D$	0.02	0.03
$\sigma_t^S$	0.25	0.5
$\text{corr}(dc/c, dS/S)$	0.2	0.4

Table 1.3: Equity Premium: parameters for the regime switching economy

transition matrix is

$$e^A = e^{tA} = \begin{pmatrix} 0.2 & 0.7 \\ 0.8 & 0.3 \end{pmatrix}.$$

Suppose  $X_0 = (1, 0)'$ . The parameters for the regimes are as in Table 1.3.

Then the expected annual excess rate of return is  $0.005 \times 0.2 + 0.03 \times 0.8 = 2.5\%$ . This improves the previous expected annual excess rate of return 0.4% by an order of magnitude, and is much closer to the true annual excess rate of return 8%.

Intuitively, the improvement is due to the fact that the representative agent (consumer) now requires additional risk premium for the regime shifts. With the additional risk premium, the model-implied annual excess rate of return is naturally higher. With more states, the Lucas economy with regime switching can match the true annual excess rate of return 8% even better. The economic intuition is that with more states, the risk premium should be higher, and then the model-implied expected annual excess rate of return should be higher.

## 1.7 Endogenous Interest Rate

### 1.7.1 General Utility

In this section, we determine the endogenous interest rate  $r$  in equilibrium. Recall that the first order conditions give  $\partial\mathcal{L}/\partial c_t = 0$  and  $\partial\mathcal{L}/\partial\lambda$ , which imply

$$\begin{aligned} e^{-\int_0^t \rho_s ds} u'(c_t) &= \lambda \zeta_t, \\ \mathbb{E} \left[ \zeta_T X_T - \int_0^T \zeta_t (e_t - c_t) dt \right] &= X_0. \end{aligned} \tag{1.15}$$

We can derive  $r_t$  and  $\theta_t$  in the following theorem.



**Theorem 1.7.1** *The endogenous interest rate  $r_t$ , and market price of risk  $\theta_t$  are given by*

$$\begin{aligned} \rho_t - r_t &= \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) \left( D_t \cdot (\mu^D(t)) \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \\ &\quad + \frac{\frac{1}{2} u''' \left( e_t + \sum_{i=1}^N D_t^i \right) \left( (D_t) \sigma^D(t) (\sigma^D(t))^\top (D_t)^\top + (\sigma_t^e e_t)^2 \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}, \\ (\theta_t)^\top &= (\theta_t^1, \theta_t^2, \dots, \theta_t^N) \\ &= \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) \left( D_t \cdot (\sigma^D(t)) \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}. \end{aligned}$$

*Proof.* See Appendix. □

**Corollary 1.7.2** *If  $N = 1$ , then*

$$\begin{aligned} \rho_t - r_t &= \frac{u''(e_t + D_t)(e_t \mu_t^e + \mu_t^D D_t) + \frac{1}{2} u'''(e_t + D_t) \left( (\sigma_t^D D_t)^2 + (\sigma_t^e e_t)^2 \right)}{u'(e_t + D_t)}, \\ \theta_t &= - \frac{u''(e_t + D_t)(\sigma_t^D D_t)}{u'(e_t + D_t)}. \end{aligned}$$

## 1.7.2 Bond Price and Yield

In this section, we determine the bond price. Denote by  $B(t, T)$  the price at time  $t$  of a zero-coupon bond with maturity time  $T$ .

**Proposition 1.7.3** *The price at time  $t$  of a time- $T$  zero-coupon bond is*

$$B(t, T) = \mathbb{E}_t \left[ e^{-\int_t^T \rho_u du} \frac{u' \left( e_T + \sum_{i=1}^N D_T^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \cdot 1 \right]$$

*Proof.* From the Euler equation (1.9), we know that the stochastic discount factor is

$$e^{-\int_t^s \rho_u du} \frac{u'(c_s)}{u'(c_t)} = e^{-\int_t^s \rho_u du} \frac{u' \left( e_s + \sum_{i=1}^N D_s^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}.$$

Using the stochastic discount factor, we can price any dividend stream. Since a zero-coupon bond pays off 1 at time  $T$ , we see that

$$B(t, T) = \mathbb{E}_t \left[ e^{-\int_t^T \rho_u du} \frac{u'(c_T)}{u'(c_t)} \cdot 1 \right]$$

$$= \mathbb{E}_t \left[ e^{-\int_t^T \rho_u du} \frac{u' \left( e_T + \sum_{i=1}^N D_T^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \cdot 1 \right].$$

□

**Remark.** As  $u'(c_t) = \lambda e^{\int_0^t \rho_u du} \zeta_t$ , we see that the bond price is also

$$\begin{aligned} B(t, T) &= \mathbb{E}_t \left[ e^{-\int_t^T \rho_u du} \frac{u'(c_T)}{u'(c_t)} \cdot 1 \right] \\ &= \mathbb{E}_t \left[ \frac{\zeta_T}{\zeta_t} \cdot 1 \right] \\ &= \mathbb{E}_t \left[ e^{-\int_t^T r_u du} \frac{Z_T}{Z_t} \cdot 1 \right] \\ &= \tilde{\mathbb{E}}_t \left[ e^{-\int_t^T r_u du} \cdot 1 \right]. \end{aligned}$$

More generally, for any contingent claim  $C$  (can be random), the price of  $C$  at time  $t$  is

$$\begin{aligned} P_t &= \mathbb{E}_t \left[ e^{-\int_t^T \rho_u du} \frac{u'(c_T)}{u'(c_t)} \cdot C \right] \\ &= \mathbb{E}_t \left[ \frac{\zeta_T}{\zeta_t} \cdot C \right] \\ &= \mathbb{E}_t \left[ e^{-\int_t^T r_u du} \frac{Z_T}{Z_t} \cdot C \right] \\ &= \tilde{\mathbb{E}}_t \left[ e^{-\int_t^T r_u du} \cdot C \right]. \end{aligned}$$

Therefore, in this sense, the probability measure  $\tilde{\mathbb{P}}$  serves as the risk-neutral measure in our general equilibrium pricing framework.

Now we consider the yield of the bonds. The *yield*  $y(t, T)$  of a zero-coupon bond at time  $t$  is defined by

$$B(t, T) e^{y(t, T)(T-t)} = 1.$$

**Proposition 1.7.4** *The yield of a time- $t$  zero-coupon bond with maturity time  $T$  is*

$$y(t, T) = -\frac{1}{T-t} \log \mathbb{E}_t \left[ e^{-\rho(T-t)} \frac{u' \left( e_T + \sum_{i=1}^N D_T^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \cdot 1 \right].$$

*Proof.* By definition, the yield of a zero-coupon bond at time  $t$  is

$$B(t, T) = e^{-y(t, T)(T-t)} \cdot 1.$$

Thus the yield is

$$\begin{aligned} y(t, T) &= -\frac{\log B(t, T)}{T-t} \\ &= -\frac{1}{T-t} \log \mathbb{E}_t \left[ e^{-\rho(T-t)} \frac{u' \left( e_T + \sum_{i=1}^N D_T^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \cdot 1 \right]. \end{aligned}$$

□

### 1.7.3 Constant Relative Risk Aversion Utility

In this section, we derive the endogenous interest rate, when the agent's utility is the Constant Relative Risk Aversion (CRRA) utility, i.e.,  $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ . For  $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ , we have

$$\begin{aligned} u'(c_t) &= u' \left( e_t + \sum_{i=1}^N D_t^i \right) = (c_t)^{-\gamma} = \left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma}, \\ u''(c_t) &= u'' \left( e_t + \sum_{i=1}^N D_t^i \right) = -\gamma \left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma-1}, \\ u'''(c_t) &= u''' \left( e_t + \sum_{i=1}^N D_t^i \right) = \gamma(\gamma+1) \left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma-2} \end{aligned}$$

Thus, we have the following proposition.

**Proposition 1.7.5** *Suppose the agent's utility is  $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ . Then*

$$\begin{aligned} \rho_t - r_t &= \frac{-\gamma \left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma-1} \left( e_t \mu_t^e + D_t \cdot \mu^D(t) \right)}{\left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma}}, \\ &+ \frac{\frac{1}{2} \gamma(\gamma+1) \left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma-2} \left( (D_t) \sigma^D(t) (\sigma^D(t))^\top (D_t)^\top + (\sigma_t^e e_t)^2 \right)}{\left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma}}, \end{aligned}$$

$$(\theta_t)^\top = (\theta_t^1, \theta_t^2, \dots, \theta_t^N)$$

$$= \frac{-\gamma \left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma-1} (D_t \cdot (\sigma^D(t)))}{\left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma}}.$$

When  $N = 1$  and  $e_t \equiv 0$ , the endogenous interest rate  $r_t$  and market price of risk  $\theta_t$  can be calculated by particularly simple expressions. Note that in the following corollary, the coefficients switch with the state process  $Y$ .

**Corollary 1.7.6** *When  $N = 1$  and  $e_t \equiv 0$ , we have*

$$\begin{aligned} \rho_t - r_t &= -\gamma \mu_t^D + \frac{1}{2} \gamma (\gamma + 1) (\sigma_t^D)^2, \\ \theta_t &= \gamma \sigma_t^D. \end{aligned}$$

*Proof.* See Appendix □

**Proposition 1.7.7** *Suppose the agent's utility is CRRA, i.e.,  $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ . Then the bond price is*

$$B(t, T) = \mathbb{E}_t \left[ e^{-\rho(T-t)} \frac{\left( e_T + \sum_{i=1}^N D_T^i \right)^{-\gamma}}{\left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma}} \cdot 1 \right]$$

*The yield of the bond is*

$$y(t, T) = -\frac{1}{T-t} \log \mathbb{E}_t \left[ e^{-\rho(T-t)} \frac{\left( e_T + \sum_{i=1}^N D_T^i \right)^{-\gamma}}{\left( e_t + \sum_{i=1}^N D_t^i \right)^{-\gamma}} \cdot 1 \right].$$

## 1.8 Term Structure of Interest Rates

Term structure of interest rates has been an important topic in the asset pricing theory. The classic contribution in Cox et al. (1985a, 1985b) derives the celebrated CIR model from a general equilibrium **production** economy.<sup>8</sup> In Amisano and Tristani (2010), and Chib,

---

<sup>8</sup>A production economy is, in a sense, the polar opposite of an endowment economy, which is considered in this paper. Roughly speaking, a production economy allows for an aggregate consumption savings, which can be invested to the production technologies. The rate of return on the production is assumed to be exogenously given. For excellent comparison and the distinction between endowment and production economies, see e.g. Pennacchi (2007).

Kang and Ramamurthy (2010), term structures of interest rates in a discrete time, dynamic stochastic general equilibrium (DSGE) model are considered. These papers then examine the Euler equation derived from the utility maximization problem, and approximate solutions for the bond prices are derived.<sup>9</sup>

The term structures in this paper differ from the previous contributions in a number of ways. Firstly, the term structure of interest rates in our endowment economy is more tractable. We are able to obtain closed-form exact solutions for bond prices.<sup>10</sup> Secondly, our endowment economy implies a rich framework for the term structure of interest rates. For example, we shall show that our endowment economy gives rise to well-known term structure models such as square-root model, three-halves model, and Gaussian models with regime switching, to name just a few important ones. Our framework also implies multi-factor affine models, as well as quadratic models.

We now investigate the rich framework for the term structure of interest rates in our endowment economy. For simplicity, in this section, take  $\gamma = 1$ . Thus the agent's utility is log utility, i.e.,  $u(c_t) = \log c_t$ .

We examine the simplest case, where  $N = 1$  and  $e_t \equiv 0$ . Then, from Corollary 1.7.6, we have

$$\begin{aligned}\rho_t - r_t &= -\gamma\mu_t^D + \frac{1}{2}\gamma(\gamma + 1)(\sigma_t^D)^2 \\ &= -\mu_t^D + (\sigma_t^D)^2, \\ \theta_t &= \sigma_t^D.\end{aligned}$$

Suppose that  $\mu_t^D$  and  $\sigma_t^D$  are given by

$$\mu_t^D = \mu_t x_t - \rho_t$$

---

<sup>9</sup>In Amisano and Tristani (2010), the authors use second-order approximation to obtain the bond prices. In Chib, Kang and Ramamurthy (2010), first-order approximation is used to obtain equilibrium conditions, and log-linear approximation is used to specify the pricing-kernel and then derive the bond prices.

<sup>10</sup>In a DSGE model, the income, labor, and wages are all endogenously determined. Thus the DSGE model is sophisticated and provides rich economic implications. Due to its complexity, people usually have to resort to approximate solutions or numerical methods to obtain the prices.

$$\sigma_t^D = \sigma_t \sqrt{x_t},$$

where  $\mu_t = \mu(Y_t)$  and  $\sigma_t = \sigma(Y_t)$  switch with the states of the Markov chain  $Y$ . Then,

$$r_t = (\mu_t - \sigma_t^2)x_t.$$

Consider the following general specification for the state variable  $x_t$  (risk factor), which gives rise to a rich and tractable term structure of interest rates. Suppose  $x_t$  is a new state variable such that

$$dx_t = (c_0 + c_1x_t^\alpha + c_2x_t^\beta)dt + c_3x_t^\eta dW_t.$$

Then

$$\begin{aligned} dr_t &= (\mu_t - \sigma_t^2)dx_t \\ &= (\mu_t - \sigma_t^2)((c_0 + c_1x_t^\alpha + c_2x_t^\beta)dt + c_3x_t^\eta dW_t), \end{aligned} \tag{1.16}$$

This general parametric specification gives rise to a number of models of term structure of interest rate with regime switching. For illustrative purposes, we only list a few important ones.

**1. Square-root model with regime switching.** Setting  $\alpha = 1$ ,  $\beta = 0$ ,  $\eta = 0.5$  in equation (1.16), we have

$$dr_t = \kappa_t(\theta_t - r_t)dt + \sigma_t\sqrt{r_t}dW_t,$$

where  $\kappa_t = \langle \kappa, Y_t \rangle$ ,  $\theta_t = \langle \theta, Y_t \rangle$ ,  $\sigma_t = \langle \sigma, Y_t \rangle$  are switching with the Markov chain  $Y$ . This is an extension of the term structure model proposed in Cox et al. (1985).

**2. Three-halves model with regime switching.** Setting  $C_0 = 0$ ,  $\alpha = 1$ ,  $\beta = 2$ ,  $\eta = 1.5$  in equation (1.16), we have

$$dr_t = \kappa_t(\theta_t - r_t)r_tdt + \sigma_t r_t^{1.5}dW_t,$$

where  $\kappa_t = \langle \kappa, Y_t \rangle$ ,  $\theta_t = \langle \theta, Y_t \rangle$ ,  $\sigma_t = \langle \sigma, Y_t \rangle$  are switching with the Markov chain  $Y$ . This is an extension of the term structure model proposed in Cox et al. (1980).

**3. Gaussian model with regime switching.** Setting  $\alpha = 1$ ,  $\beta = 0$ ,  $\eta = 0$  in equation (1.16), we have

$$dr_t = \kappa_t(\theta_t - r_t)dt + \sigma_t dW_t,$$

where  $\kappa_t = \langle \kappa, Y_t \rangle$ ,  $\theta_t = \langle \theta, Y_t \rangle$ ,  $\sigma_t = \langle \sigma, Y_t \rangle$  are switching with the Markov chain  $Y$ . This is an extension of the term structure model proposed in Vasicek (1977).

**Remark.** For simplicity, we only list one-factor interest rate models here. The above approach can be easily extended to obtain multi-factor interest rate models with regime switching coefficients.

The dynamics of the short (instantaneous) rate  $r_t$  gives the affine term structure of the interest rates.

**Proposition 1.8.1** *Under the above assumptions, the price at time  $t$  of a zero-coupon bond with maturity  $T$  is*

$$B(t, T) = \exp(G(t, T)r + H(t, T)),$$

for some deterministic functions  $G$  and  $H$  of time.

*Proof.* See Appendix. □

**Remark.** In this section, for simplicity, we have only discussed a one-factor affine term structure of interest rates implied by our economy. With multi-factor specification, our model is able to give rise to a more complicated term structure of interest rates, for example, an affine term structure with many factors. Also, our equilibrium economy can give a non-affine term structure of interest rates, for example, a quadratic term structure of interest rates.

## 1.9 Conclusion

In this paper, we considered a Lucas economy in continuous time, with multiple dividend streams and regime switching.

We determined the asset prices in equilibrium in the economy, and derived a system of partial differential equations for the asset prices and the short interest rate.

We obtained the solution for the endogenous short interest rate, the bond price, and the yield of the bond. We also considered the application of the equilibrium model and showed that the model implies a rich framework for the term structure of interest rates.

In this paper, for simplicity, we only presented a one-factor affine term structure of interest rates as implied by our economy. With multi-factor specification, our model is able to give rise to a more complicated term structure of interest rates, for example, an affine term structure with many factors. Also, our equilibrium economy can provide a non-affine term structure of interest rates, for example, a quadratic term structure of interest rates.

Instead of stochastic dynamic programming as used in previous work, we applied the martingale/duality approach as developed in Karatzas, Lehocky and Shreve (1987), Cox and Huang (1989) and Pliska (1986).

## Appendix

*Proof of Equation (1.3).* We note that

$$\begin{aligned} dX_t &= d\left(\frac{X_t}{S_t^0} S_t^0\right) = \frac{X_t}{S_t^0} dS_t^0 + d\left(\frac{X_t}{S_t^0}\right) S_t^0 \\ &= \frac{X_t}{S_t^0} r_t S_t^0 dt + d\left(\frac{X_t}{S_t^0}\right) S_t^0 \\ &= r_t X_t dt + d\left(\frac{X_t}{S_t^0}\right) S_t^0. \end{aligned}$$

Comparing this equality with the budget constraint, we see that

$$\begin{aligned} dX_t &= dI_t + r_t X_t dt + \pi_t^1 ((\mu_t - r_t) dt + \sigma_t dW_t) \\ &= r_t X_t dt + d\left(\frac{X_t}{S_t^0}\right) S_t^0, \end{aligned}$$

and hence

$$d\left(\frac{X_t}{S_t^0}\right) S_t^0 = dI_t + \pi_t^1 ((\mu_t - r_t) dt + \sigma_t dW_t).$$



Thus

$$d\left(\frac{X_t}{S_t^0}\right) = \frac{\pi_t^1((\mu_t - r_t)dt + \sigma_t dW_t) + dI_t}{S_t^0}.$$

□

*Proof of Proposition 1.4.3.* From equation (1.9), with  $u(C_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ , we have

$$\begin{aligned} S_t^i &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \frac{u'(c_s)}{u'(c_t)} D_s^i ds \right] \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(\frac{c_s}{c_t}\right)^{-\gamma} D_s^i ds \right] \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(\frac{e_s + \sum_{i=1}^N D_s^i}{e_t + \sum_{i=1}^N D_t^i}\right)^{-\gamma} D_s^i ds \right] \quad (c_s = e_s + \sum_{i=1}^N D_s^i) \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(\frac{e_s + \sum_{i=1}^N D_s^i}{e_t + \sum_{i=1}^N D_t^i}\right)^{-\gamma} D_t^i \exp\left(\int_t^s \mu_i^D(u) - \frac{1}{2}\|\sigma_i^D(u)\|^2 du + \int_t^s \sigma_i^D(u) \cdot dW_u\right) ds \right] \\ &= D_t^i \mathbb{E}_t \left[ \int_t^T \left(\frac{e_s + \sum_{i=1}^N D_s^i}{e_t + \sum_{i=1}^N D_t^i}\right)^{-\gamma} \exp\left(\int_t^s -\rho_u du\right) \exp\left(\int_t^s \mu_i^D(u) - \frac{1}{2}\|\sigma_i^D(u)\|^2 du + \int_t^s \sigma_i^D(u) \cdot dW_u\right) ds \right] \end{aligned}$$

□

*Proof of Theorem 1.4.4.* Again, the standard Euler equation of the pricing formula is

$$e^{-\int_0^t \rho_s ds} u'(c_t) S_t = \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} u'(c_s) D_s ds \right].$$

Thus, using Proposition 1.4.2, with  $u(C_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$ ,

$$\begin{aligned} S_t &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \frac{u'(c_s)}{u'(c_t)} D_s ds \right] \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(\frac{c_s}{c_t}\right)^{-\gamma} D_s ds \right] \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(\frac{D_s}{D_t}\right)^{-\gamma} D_s ds \right] \quad (c_s = D_s) \\ &= \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(\frac{D_s}{D_t}\right)^{1-\gamma} D_t ds \right] \\ &= D_t \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} \left(e^{\int_t^s \mu_u^D - \frac{1}{2}(\sigma_u^D)^2 du + \int_t^s \sigma_u^D dW_u}\right)^{1-\gamma} ds \right] \end{aligned}$$

$$= D_t \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} e^{(1-\gamma)(\int_t^s \mu_u^D - \frac{1}{2}(\sigma_u^D)^2 du + \int_t^s \sigma_u^D dW_u)} ds \right].$$

Now

$$\mathbb{E} \left[ \int_t^T \exp \left( \int_t^u \sigma_s^D dW_s \right) du \middle| \mathcal{F}_t, \mathcal{F}_t^Y \right] = \mathbb{E} \left[ \int_t^T \exp \left( \int_t^u \frac{1}{2} (\sigma_s^D)^2 ds \right) du \middle| \mathcal{F}_t, \mathcal{F}_t^Y \right],$$

which implies

$$\mathbb{E} \left[ \int_t^T \exp \left( \int_t^u (1-\gamma) \sigma_s^D dW_s \right) du \middle| \mathcal{F}_t, \mathcal{F}_t^Y \right] = \mathbb{E} \left[ \int_t^T \exp \left( \int_t^u \frac{1}{2} ((1-\gamma) \sigma_s^D)^2 ds \right) du \middle| \mathcal{F}_t, \mathcal{F}_t^Y \right].$$

Thus

$$\begin{aligned} S_t &= D_t \mathbb{E}_t \left[ \int_t^T e^{-\int_t^s \rho_u du} e^{(1-\gamma)(\int_t^s \mu_u^D - \frac{1}{2}(\sigma_u^D)^2 + \int_t^s \frac{1}{2}((1-\gamma)\sigma_u^D)^2 du)} ds \right] \\ &= D_t \mathbb{E} \left[ \int_t^T \exp \left( \int_t^s (-\rho_u + (1-\gamma)\mu_u^D - \frac{1}{2}(1-\gamma)(\sigma_u^D)^2 + \frac{1}{2}(1-\gamma)^2(\sigma_u^D)^2) du \right) ds \middle| \mathcal{F}_t \right] \\ &= D_t \mathbb{E} \left[ \int_t^T \exp \left( \int_t^s (-\rho_u + (1-\gamma)\mu_u^D - \frac{1}{2}\gamma(1-\gamma)(\sigma_u^D)^2) du \right) ds \middle| \mathcal{F}_t \right]. \end{aligned}$$

Write

$$\lambda_u = -\rho_u + (1-\gamma)\mu_u^D - \frac{1}{2}\gamma(1-\gamma)(\sigma_u^D)^2.$$

Then

$$S_t = D_t \mathbb{E} \left[ \int_t^T \exp \left( \int_t^s \lambda_u du \right) ds \middle| \mathcal{F}_t, \mathcal{F}_t^Y \right].$$

As  $\sigma_u^D$ ,  $\mu_u^D$ ,  $\lambda_u$  switch according to the state of  $Y$ , we take the following approach to calculate  $S_t$ . To find

$$S_t = D_t \mathbb{E}_t \left[ \int_t^T \exp \left( \int_t^s \lambda_u du \right) ds \right],$$

we need to obtain

$$\mathbb{E}_t \left[ \int_t^T \exp \left( \int_t^s \lambda_u du \right) ds \right].$$

For a fixed  $t > 0$ , define

$$\Gamma_s := \exp \left( \int_t^s \lambda_u du \right) \cdot Y_s \in \mathbb{R}^K.$$

Then, using equation (1.2)

$$d\Gamma_s = \lambda_s \Gamma_s ds + \exp\left(\int_t^s \lambda_u du\right) (AY_s du + dM_s),$$

and hence

$$\int_t^s d\Gamma_u = \int_t^s \lambda_u \Gamma_u du + \exp\left(\int_t^s \lambda_u du\right) (AY_u du + dM_u).$$

Consequently,

$$\begin{aligned} \mathbb{E}_t[\Gamma_s] - \Gamma_t &= \int_t^s \mathbb{E}_t[\lambda_u \Gamma_u] du + \mathbb{E}_t\left[\exp\left(\int_t^s \lambda_u du\right) (AY_u du + dM_u)\right] \\ &= \int_t^s (A + \text{diag } \lambda) \mathbb{E}_t[\Gamma_u] du. \end{aligned}$$

Thus

$$\mathbb{E}_t[\Gamma_s] = \Gamma_t + \int_t^s (A + \text{diag } \lambda) \mathbb{E}_t[\Gamma_u] du.$$

Hence

$$\mathbb{E}_t\left[\exp\left(\int_t^s \lambda_u du\right)\right] = \langle \mathbb{E}_t[\Gamma_s], \mathbf{1} \rangle,$$

and

$$\mathbb{E}_t\left[\int_t^T \exp\left(\int_t^s \lambda_u du\right) ds\right] = \int_t^T \langle \mathbb{E}_t[\Gamma_s], \mathbf{1} \rangle ds.$$

Therefore

$$S_t = D_t \int_t^T \langle \mathbb{E}_t[\Gamma_s], \mathbf{1} \rangle ds.$$

Now, for fixed  $t \geq 0$ , we solve for  $\mathbb{E}_t[\Gamma_s]$  from the linear system of ordinary differential equations

$$\mathbb{E}_t[\Gamma_s] = \Gamma_t + \int_t^s (A + \text{diag } \lambda) \mathbb{E}_t[\Gamma_u] du.$$

The solution of the system of ordinary differential equations is

$$\mathbb{E}_t[\Gamma_s] = e^{(A + \text{diag } \lambda)(s-t)} Y_t.$$

Thus

$$S_t = D_t \int_t^T \langle e^{(A + \text{diag } \lambda)(s-t)} Y_t, \mathbf{1} \rangle ds.$$

□

*Proof of Theorem 1.5.1.* Suppose  $S_t^i = f_i(D_t)$ . Then, Ito's lemma gives

$$\begin{aligned} dS_t^i &= f'_i(D_t^i) (D_t^i \mu_i^D(t) dt + D_t^i \sigma_i^D(t) dW_t) + \frac{1}{2} f''_i(D_t^i) (D_t^i)^2 \|\sigma_i^D(t)\|^2 dt \\ &= \left( f'_i(D_t^i) D_t^i \mu_i^D(t) + \frac{1}{2} f''_i(D_t^i) (D_t^i)^2 \|\sigma_i^D(t)\|^2 \right) dt + f'_i(D_t^i) D_t^i \sigma_i^D(t) dW_t. \end{aligned}$$

Suppose there exist  $\mu_i^S(t)$  and  $\sigma_i^S(t)$  such that

$$dS_t^i = S_t^i (\mu_i^S(t) dt + \sigma_i^S(t) dW_t).$$

Comparing the coefficients in the equations, we have

$$\begin{aligned} \mu_i^S(t) &= \frac{f'_i(D_t^i) D_t^i \mu_i^D(t) + \frac{1}{2} f''_i(D_t^i) (D_t^i)^2 \|\sigma_i^D(t)\|^2}{f_i(D_t)}, \\ \sigma_i^S(t) &= \frac{f'_i(D_t^i) D_t^i \sigma_i^D(t)}{f_i(D_t)}. \end{aligned}$$

Also  $\theta_t = (\theta_t^1, \dots, \theta_t^N)^\top$  is such that

$$\alpha_i^S(t) - r_t = \sum_{j=1}^N \sigma_{ij}^S(t) \theta_t^j, \quad i = 1, \dots, m,$$

or equivalently,

$$\theta = (\sigma_t)^{-1} (\alpha^S(t) - r_t \cdot \mathbf{1}).$$

Thus we have

$$\sum_{j=1}^N \sigma_{ij}^S(t) \theta_t^j = \frac{f'_i(D_t^i) D_t^i \mu_i^D(t) + \frac{1}{2} f''_i(D_t^i) (D_t^i)^2 \|\sigma_i^D(t)\|^2}{f_i(D_t)} - r_t + \frac{D_t^i}{f_i(D_t^i)},$$

and hence

$$r_t f_i(D_t^i) = D_t^i + D_t^i \mu_i^D(t) f'_i(D_t^i) - \left( \sum_{j=1}^N \sigma_{ij}^S(t) \theta_t^j \right) f_i(D_t^i) + \frac{1}{2} f''_i(D_t^i) (D_t^i)^2 \|\sigma_i^D(t)\|^2,$$

for  $i = 1, 2, \dots, N$ . □

*Proof of Theorem 1.7.1.* Since the agent maximizes the utility only from consumption  $c_t$ , we know  $X_T = 0$ . From the equations in (1.15), we see that

$$u'(c_t) = \lambda e^{\int_0^t \rho_s ds} \zeta_t,$$

and  $\lambda = u'(c_0)$ . Since in equilibrium,  $c_t = \sum_{i=1}^N D_t^i$ , we have

$$\begin{aligned} u' \left( e_t + \sum_{i=1}^N D_t^i \right) &= u'(c_t) = \lambda e^{\int_0^t \rho_s ds} \zeta_t \\ &= \lambda \exp \left( \int_0^t \rho_u du - \int_0^t r_u du \right) \exp \left( - \int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du \right). \end{aligned}$$

The dynamics of  $D_t$  and  $e_t$  are exogenously given, so we can obtain  $r_t$  from the above equation. From the left-hand side of the equation, Ito's lemma gives

$$\begin{aligned} du' \left( e_t + \sum_{i=1}^N D_t^i \right) &= u'' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) + \frac{1}{2} u''' \left( e_t + \sum_{i=1}^N D_t^i \right) d \langle e_t + \sum_{i=1}^N D_t^i, e_t + \sum_{i=1}^N D_t^i \rangle \\ &= u'' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) + \frac{1}{2} u''' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) \end{aligned}$$

From the right-hand side of the equation, Ito's lemma gives

$$\begin{aligned} & d \exp \left( \int_0^t \rho_u du - \int_0^t r_u du - \int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du \right) \\ &= \exp \left( \int_0^t \rho_u du - \int_0^t r_u du - \int_0^t \theta_u dW_u - \frac{1}{2} \int_0^t \theta_u^2 du \right) \left( \rho_t dt - r_t dt - \theta_t dW_t - \frac{1}{2} \theta_t^2 dt + \frac{1}{2} \theta_t^2 dt \right) \\ &= u' \left( e_t + \sum_{i=1}^N D_t^i \right) (\rho_t dt - r_t dt - \theta_t dW_t) \end{aligned}$$

Comparing these, we see that

$$\begin{aligned} \rho_t dt - r_t dt - \theta_t dW_t &= \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \\ &\quad + \frac{\frac{1}{2} u''' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}. \end{aligned}$$

Equating the  $dt$  and  $dW_t$  terms, we have the characterization for the endogenous interest rate  $r_t$ . Note that

$$\begin{aligned} d \left( e_t + \sum_{i=1}^N D_t^i \right) &= \sum_{i=1}^N \left( D_t^i \mu_i^D(t) dt + D_t^i \sum_{j=1}^N \sigma_{ij}^D(t) dW_t^j \right) + de_t \\ &= \sum_{i=1}^N \left( D_t^i \mu_i^D(t) dt + D_t^i (\sigma_i^D(t) \cdot dW_t) \right) + de_t \end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} D_t^1 & D_t^2 & \cdots & D_t^N \end{bmatrix} (\mu^D(t)dt + \sigma^D(t) \cdot dW_t) + de_t \\
&= D_t \cdot (\mu^D(t)dt + \sigma^D(t) \cdot dW_t) + de_t,
\end{aligned}$$

where

$$\sigma_i^D(t) = \begin{bmatrix} \sigma_{i1}^D(t) & \sigma_{i2}^D(t) & \cdots & \sigma_{iN}^D(t) \end{bmatrix}.$$

Recall that  $de_t = e_t(\mu_t^e dt + \sigma_t^e dW_t^e)$ , where  $W^e$  is independent of  $\{W^1, \dots, W^N\}$ , and hence

$dW^e \cdot dW^i = 0$  for  $i = 1, \dots, N$ . Then

$$\begin{aligned}
d \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) &= \begin{bmatrix} D_t^1 & D_t^2 & \cdots & D_t^N \end{bmatrix} (\sigma^D(t) \cdot dW_t) \left( \begin{bmatrix} D_t^1 & D_t^2 & \cdots & D_t^N \end{bmatrix} (\sigma^D(t) \cdot dW_t) \right)^\top \\
&\quad + (\sigma_t^e e_t dW_t^e) (\sigma_t^e e_t dW_t^e) \\
&= (D_t) (\sigma^D(t) \cdot dW_t) \left( D_t (\sigma^D(t) \cdot dW_t) \right)^\top + (\sigma_t^e e_t dW_t^e) (\sigma_t^e e_t dW_t^e) \\
&= (D_t) \sigma^D(t) \cdot dW_t (dW_t)^\top (\sigma^D(t))^\top (D_t)^\top + (\sigma_t^e e_t)^2 dt \\
&= (D_t) \sigma^D(t) dt (\sigma^D(t))^\top (D_t)^\top + (\sigma_t^e e_t)^2 dt \\
&= (D_t) \sigma^D(t) (\sigma^D(t))^\top (D_t)^\top dt + (\sigma_t^e e_t)^2 dt.
\end{aligned}$$

Then

$$\begin{aligned}
\rho_t dt - r_t dt - \theta_t dW_t &= \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \\
&\quad + \frac{\frac{1}{2} u''' \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right) d \left( e_t + \sum_{i=1}^N D_t^i \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \\
&= \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) (de_t + D_t \cdot (\mu^D(t)dt + \sigma^D(t) \cdot dW_t))}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)} \\
&\quad + \frac{\frac{1}{2} u''' \left( e_t + \sum_{i=1}^N D_t^i \right) ((D_t) \sigma^D(t) (\sigma^D(t))^\top (D_t)^\top dt + (\sigma_t^e e_t)^2 dt)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}.
\end{aligned}$$

Equating the  $dt$  and  $dW_t$  terms, we have

$$\rho_t - r_t = \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) \left( e_t \mu_t^e + D_t \cdot \mu^D(t) \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}$$

$$\begin{aligned}
& + \frac{\frac{1}{2}u''' \left( e_t + \sum_{i=1}^N D_t^i \right) \left( (D_t)\sigma^D(t)(\sigma^D(t))^\top (D_t)^\top + (\sigma_t^e e_t)^2 \right)}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}, \\
(\theta_t)^\top &= (\theta_t^1, \theta_t^2, \dots, \theta_t^N) \\
&= \frac{u'' \left( e_t + \sum_{i=1}^N D_t^i \right) (D_t \cdot (\sigma^D(t)))}{u' \left( e_t + \sum_{i=1}^N D_t^i \right)}.
\end{aligned}$$

□

*Proof of Corollary 1.7.6.* If  $N = 1$  and  $e_t \equiv 0$ , from Corollary 1.7.2 we have

$$\begin{aligned}
\rho_t - r_t &= \frac{u''(D_t)(\mu_t^D D_t) + \frac{1}{2}u'''(D_t)(\sigma_t^D D_t)^2}{u'(D_t)} \\
&= \frac{-\gamma(D_t)^{-\gamma-1}(\mu_t^D D_t) + \frac{1}{2}\gamma(\gamma+1)(D_t)^{-\gamma-2}(\sigma_t^D D_t)^2}{(D_t)^{-\gamma}} \\
&= -\gamma \frac{\mu_t^D D_t}{D_t} + \frac{1}{2}\gamma(\gamma+1) \frac{(\sigma_t^D D_t)^2}{(D_t)^2} \\
&= -\gamma \mu_t^D + \frac{1}{2}\gamma(\gamma+1)(\sigma_t^D)^2, \\
\theta_t &= -\frac{u''(D_t)(D_t \sigma_t^D)}{u'(D_t)} \\
&= -\frac{-\gamma(D_t)^{-\gamma-1}(D_t \sigma_t^D)}{(D_t)^{-\gamma}} \\
&= \gamma \frac{D_t \sigma_t^D}{D_t} \\
&= \gamma \sigma_t^D.
\end{aligned}$$

□

*Proof of Proposition 1.8.1.* The price of the bond satisfies the coupled Feynman-Kac type equation

$$V_t + \frac{1}{2}\sigma^2 V_{rr} + \mu V_r - rV = 0, \quad V(T, r) = 1. \quad (1.17)$$

We have

$$dr_t = \kappa(Y_t)(\bar{r}(Y_t) - r_t)dt + \tilde{\beta}(Y_t)\sqrt{r_t}dW_t$$

$$=: \mu(t, r_t)dt + \sigma(t, r_t)dW_t,$$

where

$$\mu(t, r_t) := \kappa(Y_t)(\bar{r}(Y_t) - r_t)$$

$$\sigma(t, r_t) := \tilde{\beta}(Y_t)\sqrt{r_t}.$$

Let  $\alpha(t) = \alpha(Y_t)$ ,  $\gamma(Y_t) = \gamma(t)$ ,  $\delta(Y_t) = \delta(t)$  be the time-dependent coefficients, switching with the state of the Markov chain  $Y$ , such that

$$\mu(t, r_t) = \alpha(Y_t)r_t + \delta(Y_t)$$

$$\sigma^2(t, r_t) = \gamma(Y_t)r_t.$$

We substitute  $\mu(t, r_t)$  and  $\sigma^2(t, r_t)$  into the Feynman-Kac equation (1.17), and replace  $V$  with  $B(t, r)$ . The equation implies

$$H_t + \delta G - r \left( 1 - G_t + \alpha G - \frac{1}{2} \gamma G^2 \right) = 0.$$

Thus we get the following two ordinary differential equations

$$1 - G_t + \alpha G - \frac{1}{2} \gamma G^2 = 0$$

$$H_t + \delta G = 0.$$

The first ODE has a unique solution satisfying  $G(T, T) = 0$ . Once we have  $G(t, T)$ , we can obtain  $H$  from the second equation as

$$H(t, T) = \int_0^T \delta(u)G(u, T)du,$$

with  $H(T, T) = 0$ . □



## Chapter 2

# Credit Risk with Latent Contagion and Frailty: Default Probabilities, Pricing, and Hedging

### 2.1 Introduction

Modeling credit risk is one of the central topics in financial research. There are generally two major approaches to model credit risks: the structural model and the reduced-form model. The structural approach usually considers the relationship between a firm's asset value and the firm's default event. The default of the firm is triggered when the asset value falls below a default barrier level. On the other hand, the reduced-form approach, introduced by Duffie and Singleton (1999), Jarrow and Turnbull (1995), treats defaults as exogenous events and models arrivals of defaults by Poisson point processes.

Credit risk models with incomplete information have been considered by various researchers. Recent literature for both structural and reduced-form models includes Collin-Dufresne, Goldstein and Helwege (2003), Schönbucher (2003), Duffie and Lando (2001), Duffie, Eckner, Horel and Saita (2009), Frey and Runggaldier (2010), and Frey and Schmidt (2012), etc. The reduced-form models in these papers accommodate some common features: for example, default intensities are driven by an underlying (state) process  $X$ ; conditional on the current information, the default times are assumed to be random times; investors have access to partial information consisting of observations of the market, as well as observable economic covariates. Collin-Dufresne *et al.* (2003) and Schönbucher (2003) model the underlying factor  $X$  as a static random vector. Moreover, it is pointed out in Collin-Dufresne *et al.* (2003) and Schönbucher (2003) that the event that some firm has defaulted causes an update of the distribution of the factor process, and hence causes a jump in the default in-

tensity of other surviving firms. Therefore, the successive updating of the distribution of the state process in reaction to incoming default observations is likely to give rise to contagion effects. Duffie *et al.* (2009) model the state process  $X$  to be an Ornstein–Uhlenbeck process. Furthermore, the empirical analysis in Duffie *et al.* (2009) suggests that, in addition to the observable economic covariates, an unobservable process driving default intensities is needed to account for the clustering of defaults in historical data.

Recent extensive empirical research in Azizpour, Giesecke and Schenkler (2014) indicates that firms usually have joint exposure to both contagion and frailty effects, in addition to the observable macro-economic risk factors. For example, empirical evidences in Azizpour, Giesecke and Schenkler (2014) indicate that contagion effect plays a more prominent role for explaining the default clusters than the frailty factor; the in-sample fit of a model including the contagion effect outperforms a model without that effect; a model ignoring the contagion effect tends to overstate the frailty effect. Out-of-sample tests in Azizpour, Giesecke and Schenkler (2014) indicate that models with contagion and frailty clustering sources perform much better on forecasting defaults than models where firms only have exposure to the observable macro-economic risk factors; a model without the contagion effect generates excessively high and volatile forecasts, while a model without the frailty effect tends to understate the forecasts. Only a model with all three sources of clustering provides accurate and sensible out-of-sample forecasts of correlated defaults. An important example for default contagion is evidenced by the default of Lehman Brothers in 2008 and the subsequent market crash.

Our paper is mainly motivated by the empirical findings presented in Azizpour, Giesecke and Schenkler (2014) that firms are usually exposed jointly to both contagion and frailty effects, in addition to the observable macro-economic risk factors. We investigate a reduced-form model driven by multiple underlying, background state processes. For simplicity and tractability, these state processes are modeled as finite-state Markov chains. In the context

of our model, the states of the Markov Chains represent different economic environments. The switching of the states of the economy can be attributed to structural changes in macro-economic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities. In our model, the intensity of defaults and the default loss sizes may both depend on the states of the economy.

In practice, investors can observe the number of defaults and individual default loss amounts over a given time period. However, the underlying states of the economy, or the regime-switching process, are not directly observable. Therefore, we further assume that the background state processes are *Hidden Markov chains*, namely the states of the chains are not directly observable. Rather, the states are estimated through the observations of other information (e.g., default events, losses, etc.), as is often the case for the investors in the financial markets.

With the unobservable state processes setting, we are able to obtain **closed-form** solutions for the joint default probabilities. We also consider the pricing of credit risk bearing securities, and obtain the pricing formulas for general credit securities.

Our model differs from the previous contributions in the following ways, and has a number of advantages. Firstly, our models have multiple underlying state processes, representing firm-specific factors. With this approach, the contagion effects arise more naturally and explicitly. In addition to the firm-specific state processes, another process is included which models the global state which affects all the firms. Thus the frailty effect is also included explicitly. The interplay of the frailty and the default contagion effects is captured in our multiple underlying state processes. In this framework, we are able to obtain closed-form solutions for the joint default probabilities. We also obtain the pricing formulas for the credit securities. Secondly, the multiple underlying processes approach that we take is flexible and versatile. For example, in this paper, for simplicity, we assume that the underlying state

processes are all unobservable. With the multiple state processes approach, some existing state process can be easily changed to model observable factors. Similarly, new state process can be added as a new component to model observable factors. For example, there are well-documented economic data and measurements for the state of the global economy, and such data can be represented by a new, observable state process in our model. Our multiple underlying processes framework also allows us to incorporate easily other economic covariates in our model. These economic covariates can be either observable or unobservable, represented by new state processes. Our method can be easily adapted to handle such cases. Thirdly, our framework allows us to give joint estimate of the parameters from time series data on both historical defaults and securities prices (e.g., corporate bond yields). In our framework, we make clear distinction between the historical (“real world”, physical) measure  $\mathbb{P}$  and the risk-neutral measure  $\mathbb{Q}$ . We follow the principle, as emphasized in Duffie and Singleton (2003), Dai and Singleton (2003), that the pricing can be performed with the risk-neutral measure  $\mathbb{Q}$ , and the parameter estimation needs to be implemented in the historical measure  $\mathbb{P}$ .<sup>1</sup> With this approach, we can obtain joint estimate of the parameters from both historical defaults and securities prices.

The paper is organized as follows. In Section 2, we give the ideas and constituents of our models. In Section 3, we define our models in detail, and give corresponding filtering results. In Section 4, we present the **closed-form** solutions for the joint default probabilities in the real world measure  $\mathbb{P}$ . In Section 5, we obtain pricing formulas for the credit risk bearing

---

<sup>1</sup>Some previous contributions work with the risk-neutral measure  $\mathbb{Q}$  all along, including the model calibration. This may cause problematic modeling and parameter estimation for a number of reasons. Firstly, the observation process itself is in the “real-world”, and thus under the historical measure  $\mathbb{P}$ . The actual probabilities under the historical measure  $\mathbb{P}$  are usually different than the risk-neutral probabilities under measure  $\mathbb{Q}$ , due to many effects, e.g., risk premia, tax shields, liquidity, etc., as is illustrated in detail in Duffie and Singleton (2003). Secondly, whether in complete or incomplete markets, knowledge of risk-neutral probabilities is generally not enough information to assess the fit of a model to historical data, as the behavior of observed prices reflects the actual, real-world, conditional distributions that generate historical data. When the objectives are both to price securities and to assess the implications of the pricing models, it is typically necessary to specify both the actual and risk-neutral probabilities. This principle, because of its importance in model estimation and assessment, has been emphasized repeatedly in Duffie and Singleton (2003). For more in depth discussions, see, e.g., Duffie and Singleton (2003), Dai and Singleton (2003), Pan (2002).

securities, using the risk-neutral measure  $\mathbb{Q}$ . In Section 6, we investigate dynamic hedging of the securities. In Section 7, we consider parameter estimation from historical defaults, and give maximum likelihood estimates for the parameters in the historical measure  $\mathbb{P}$ , by using the EM algorithm. In Section 7 we also discuss the joint estimate of the parameters from time series data on both historical defaults and securities prices. This is achieved by a parametric specification of the pricing kernel. In Section 8, we give some concluding remarks.

## 2.2 An Ideal Model Dynamics and Methodology

In this section, we describe our “ideal” model dynamics, which allows for both *contagion* and *frailty*. We shall also briefly summarize our methodology and tools.

Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a complete, filtered probability space, with the filtration  $\mathbb{F} = (\mathcal{F})_{t \geq 0}$  satisfying the usual conditions. We suppose that the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  is rich enough to model the randomness of the observations process and of the state processes. Thus, all stochastic processes considered are, by definition,  $\mathbb{F}$ -adapted. We shall use this fact in various places in this paper.

The probability measure  $\mathbb{P}$  is assumed to be the historical (the “real world”, or physical) measure. The risk-neutral measure where the pricing of securities takes place is denoted by  $\mathbb{Q}$ . The measures  $\mathbb{P}$  and  $\mathbb{Q}$  are linked by the pricing kernel (state price density, or stochastic discount factor), which is explained in more detail in Section 2.7.2.

### 2.2.1 An Ideal Model Dynamics with Contagion and Frailty

Suppose we have  $N$  defaultable securities issued by  $N$  firms. We are concerned with the credit derivatives, i.e., derivative securities whose payoffs depends on the *default events* of the  $N$  firms.

For each firm  $i$ ,  $1 \leq i \leq N$ , there is an underlying state process  $X^i = (X_t^i)_{t \geq 0}$ . The process  $X^i$  can be thought of as representing the firm-specific state of the  $i$ th firm. In

addition, there is also a process  $X = (X_t)_{t \geq 0}$ , which represents the global state that affects all firms  $i$ ,  $1 \leq i \leq N$ . This could be the global economy or a political regime.

In the context of our model, the states of processes  $X, X^1, X^2, \dots, X^N$  represent different economic environments. The switching of the states of the economy can be attributed to structural changes in macro-economic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities.

For tractability, we suppose that the processes  $X, X^1, X^2, \dots, X^N$  are all continuous-time, finite-state Markov chains. For simplicity, we further assume the processes  $X, X^1, X^2, \dots, X^N$  are independent.

For each firm  $i$ ,  $1 \leq i \leq N$ , the default time  $\tau_i$  is assumed to be a random time, with

$$\mathbb{P}(\tau_i \geq t | \mathcal{F}_t) = \exp(-C_t^i),$$

where  $C_t^i$  is a functional of the form

$$C_t^i = \int_0^t \lambda_s^i ds.$$

Here  $\lambda_s^i$  is the default intensity of the  $i$ th firm, depending on time  $s$ . In our model, we could suppose that the default intensity  $\lambda_s^i$  is a function of all state processes  $X, X^1, X^2, \dots, X^N$ . Precisely,  $\lambda_t^i$  could be given by

$$\lambda_t^i = \lambda^i(X_t, X_t^1, X_t^2, \dots, X_t^N) = \alpha_0^i(X_t) + \alpha_1^i(X_t^1) + \alpha_2^i(X_t^2) + \dots + \alpha_N^i(X_t^N). \quad (2.1)$$

Here,  $\alpha_k^i(\cdot)$  is a function of the state of the  $k$ th chain  $X^k$ ,  $1 \leq k \leq N$ . Consequently, we would have

$$\lambda_t^1 = \lambda^1(X_t, X_t^1, X_t^2, \dots, X_t^N) = \alpha_0^1(X_t) + \alpha_1^1(X_t^1) + \alpha_2^1(X_t^2) + \dots + \alpha_N^1(X_t^N)$$

$$\lambda_t^2 = \lambda^2(X_t, X_t^1, X_t^2, \dots, X_t^N) = \alpha_0^2(X_t) + \alpha_1^2(X_t^1) + \alpha_2^2(X_t^2) + \dots + \alpha_N^2(X_t^N)$$

$$\vdots \quad \vdots \quad \vdots$$

$$\lambda_t^N = \lambda^N(X_t, X_t^1, X_t^2, \dots, X_t^N) = \alpha_0^N(X_t) + \alpha_1^N(X_t^1) + \alpha_2^N(X_t^2) + \dots + \alpha_N^N(X_t^N).$$

We note that, the processes  $X^1, X^2, \dots, X^N$  model the “*contagion*”, by which the default (or even the state switching) of one firm could have a direct influence on the default intensity of other firms. The process  $X$ , on the other hand, has the implications of “*frailty*”, by which many firms could be jointly exposed to one or more unobservable risk factors. In this sense, our model allows for both *contagion* and *frailty*.

The *dollar amount loss*, given the default of firm  $i$ , will be denoted by the random variable  $\ell_i \in (0, \infty)$ . We assume  $\ell_1, \dots, \ell_N$  are independent random variables. The *percentage loss size*  $\ell_i$  is a random variable with a density function  $f^i(\tau_i, \ell)$ . Here the parameter  $\tau$  in the function  $f^i(\tau, \ell)$  indicates that the density may depend on the default time  $\tau_i$ . For  $1 \leq i \leq N$ , define  $Z_t^i := \ell_i \mathbf{1}_{\{\tau_i \leq t\}}$ . Then the *loss state* of the credit derivatives is given by the multi-dimensional vector process  $(Z_t)_{t \geq 0} = (Z_t^1, \dots, Z_t^N)^\top$ .

### 2.2.2 Point Processes and Random Measure

The pricing of credit derivatives in our model will be modeled by random measures. In this subsection, we give a brief discussion on point processes and random measures that will be used in this paper.

Recall that for each firm  $i$ ,  $1 \leq i \leq N$ , the default time is denoted  $\tau_i$ . Then each pair

$$(\tau_i, \ell_i), \quad 1 \leq i \leq N,$$

gives a representation of the single-jump process  $(Z_t^i)_{t \geq 0}$  as a *marked point process*, with the *mark space*  $E := (0, \infty)$ . We assume  $\ell_1, \dots, \ell_N$  are random variables, with a density function  $f^i(\tau, \ell)$ . Note that when the  $i$ th firm defaults at time  $\tau_i$ , the density function  $f^i(\tau_i, \ell)$  for the random loss size depends on the random time  $\tau_i$ . Also, the density function  $f^i(\tau_i, \ell)$  depends on the states  $(X_{\tau_i-}, X_{\tau_i-}^1, X_{\tau_i-}^2, \dots, X_{\tau_i-}^N)$ .

For each  $i$ ,  $1 \leq i \leq N$ , suppose  $\mu^i(\cdot, \cdot)$  is a *random measure* on  $[0, \infty) \times E$ , which gives the random time of the  $i$ th firm’s default and the random loss size  $\ell_i \in (0, \infty)$ ,  $1 \leq i \leq N$ .

More precisely, the random measure  $\mu^i(dt, de)$  is a random delta function

$$\mu^i(dt, de) = \delta_{\tau_i}(dt)\delta_{\ell_i}(de),$$

where  $\delta_x(\cdot)$  is a Dirac delta function, or a point mass, at the point  $x$ .

For a function  $g : \Omega \times [0, \infty) \times E \rightarrow \mathbb{R}$ ,

$$\int_0^t \int_E g(\omega, u, e) \mu^i(du, de) = g(\omega, \tau_i, \ell_i) \mathbf{1}_{\{\tau_i \leq t\}}.$$

Recall

$$Z_t^i = \mathbf{1}_{\{t \geq \tau_i\}} \ell_i = \begin{cases} 0 & t < \tau_i, \\ \ell_i & t \geq \tau_i. \end{cases}$$

Then the loss of firm  $i$  at time  $t$  can be written as

$$Z_t^i = \int_0^t \int_E \ell_i \mu^i(du, de) = \ell_i \mathbf{1}_{\{\tau_i \leq t\}}.$$

Hence, the cumulative loss of all the  $N$  firms, by time  $t$ , can be written as

$$\sum_{i=1}^N \int_0^t \int_E \ell_i \mu^i(du, de).$$

In particular, let  $N_t$  denote the number of default events up to time  $t$ . Then, in terms of the random measures  $\mu^i(\cdot, \cdot)$ , the counting process  $N_t$  is

$$N_t = \sum_{i=1}^N \int_0^t \int_E \mu^i(du, de).$$

### 2.2.3 Dividend Streams

In order for us to determine the fair price of a credit derivative, the information of a *dividend stream* is needed.

Suppose there are  $N$  liquidly traded credit derivatives in the credit market. Most credit derivatives feature various kinds of intermediate cash flows,  $D_t^i$ , including payments at default dates, either to the holders or the issuers of the derivative.



## Construction of General Dividends

In general, we consider a contingent claim (credit security) which has the following features. For a detailed description, the reader is referred to Bielecki and Rutkowski (2002).

The contingent claim has par value  $Y$ , which represents the payoff at maturity, if no default occurred prior to or at time  $T$ . If default occurred prior to or at time  $T$ , the holder of the claim receives a recovery claim  $Y'$  at time  $T$ . The holder of the security may also receive a recovery payoff  $R_\tau$  at the time of default  $\tau$ . The owner of the claim may as well receive (or pay out) intermediate, cumulative cash flow (e.g., bond yields, dividends, or CDS premium payouts), which is denoted by  $(C_t)_{0 \leq t \leq T}$ . Then the cumulative dividend process  $D$  of a defaultable contingent claim can be written as

$$\begin{aligned} D_t &= (Y\mathbf{1}_{\{\tau > T\}} + Y'\mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^t (1 - N_u) dC_u + \int_0^t R_u dN_u \\ &= (Y\mathbf{1}_{\{\tau > T\}} + Y'\mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^{t \wedge \tau} dC_u + R_\tau \mathbf{1}_{\{\tau \leq t\}}. \end{aligned}$$

## Defaultable Corporate Bonds

For a corporate defaultable bond, the cumulative dividend stream  $D_t$  can be modeled as follows. Suppose the bond has par value  $Y$ . If the firm defaults, then the recovery value of the bond at time  $T$  is  $Y' = \alpha Y$ . We denote the cumulative coupon by  $(C_t)_{0 \leq t \leq T}$ . Then the bond has cumulative dividend process

$$\begin{aligned} D_t &= (Y\mathbf{1}_{\{\tau > T\}} + Y'\mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^t (1 - N_u) dC_u \\ &= (Y\mathbf{1}_{\{\tau > T\}} + \alpha Y\mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^{t \wedge \tau} dC_u. \end{aligned}$$

## Credit Default Swaps

A credit swap can be viewed as default insurance on reference loans or bonds. It pays the buyer of protection a given contingent amount at the time of a given credit event. If the insured event is a default, then the credit swap is known as a *credit default swap* (CDS). The contingent amount to be paid to the holder of a CDS at the time of the credit event

(default, say) is referred to as *default payment*. In compensation for the default payment, the CDS holder pays the issuer of the CDS an annuity, at a rate variously called the credit-swap *spread*, CDS rate, or the credit-swap *premium*. This annuity stream is paid until the maturity of the CDS, or until the time of the credit event.

Suppose the protection buyer pays the annuity (the credit-swap premium) at times  $t_i$ ,  $i = 1, 2, \dots, m$ , prior to default or maturity, whichever is earlier. For simplicity, we postulate that reference asset is a zero-coupon bond with par value  $L$  and maturity  $T$ . The default payment at time  $t$  is then

$$(L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}\mathbf{1}_{\{\tau\}}(t) = (L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}\mathbf{1}_{\{t=\tau\}},$$

where  $D(\tau)$  is the recovery value of the reference bond at the time default. On the other hand, the premium payment at time  $t$  is

$$\sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{t_i\}}(t) = \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{t=t_i\}},$$

where  $\kappa$  denotes the annuity amount. Thus, from the CDS holder's perspective, the cumulative dividend cash flow by time  $t$  is

$$\begin{aligned} D_t &= \int_0^t (L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}\mathbf{1}_{\{\tau\}}(u)du - \int_0^t \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{t_i\}}(u)du \\ &= \int_0^t (L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}dN_u - \int_0^t \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{u=t_i\}}du. \end{aligned}$$

When we postulate the fractional recovery  $\delta$  of par value, the cash flow can be further simplified as

$$D_t = \int_0^t L(1 - \delta)\mathbf{1}_{\{\tau \leq T\}}dN_u - \int_0^t \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{u=t_i\}}du.$$

### Default Put Options

A default put option is similar to a credit default swap. The difference between a default put option and a CDS is the the protection buyer pays a lump sum *premium* upfront instead

of an annuity to the issuer at the contract's inception. Therefore the cumulative cash flow by time  $t$  is

$$D_t = \int_0^t (L - D(\tau)) \mathbf{1}_{\{\tau \leq T\}} dN_u - \kappa.$$

The market price of a default option can be obtained similarly to that of a CDS.

### Other Credit Derivatives

Similarly we can derive the market price of other credit derivatives. Once the cumulative dividend stream is written out, it is relatively straightforward to obtain the market price, within our pricing framework. Other similar credit derivatives include, for example, credit linked notes (CLN), credit spread swaps, and credit spread options.

More complicated portfolio credit derivatives can be similarly priced, using the pricing equation in our model. For example, basket credit swaps,  $i$ th-to-default CDS,  $m$ -of- $n$ -to-default CDS, synthetic CDO tranche, etc. The cumulative cash flow can be written out similarly to the previous examples, with more involved analyses. The reader is referred to Bielecki and Rutkowski (2002) for more details about the cash flows for these basket credit derivative. Once the cumulative dividend stream is determined, the market price can be computed as before.

#### 2.2.4 Payoffs of Credit Derivatives

Consider a market of  $N$  credit derivatives with common maturity  $T$ . With *cumulative* dividend streams  $D_t^i$  given,  $1 \leq i \leq N$ , the discounted cumulative dividend streams are

$$\tilde{D}_t := \int_0^t e^{-\int_0^s r_u du} dD_s.$$

Then the discounted payoff of a credit derivative when it matures is (it is indeed the total sum of discounted payments in the time interval  $(t, T]$ )

$$H_{t,T}^i := \int_t^T e^{-\int_0^s r_u du} dD_s = \int_t^T d\tilde{D}_s = \tilde{D}_T^i - \tilde{D}_t^i.$$

The **full-information** filtration  $(\mathcal{F}_t)_{t \geq 0}$ ,  $\mathcal{F}_t := \sigma\{(X_s, Z_s) : s \leq t\}$ , is generated by the processes  $(X_t)_{t \geq 0}, (Z_t)_{t \geq 0}$ , up to time  $t$ . We can imagine that, for an econometrician who knows both the observation  $Z$  and the unobservable latent factor  $X$ , the value of a security with dividend stream  $D$  is the expectation of the discounted payoff  $H_{t,T}$  conditional on the full information  $\mathcal{F}_t$  up to time  $t$ , i.e.,

$$p_{t,T}^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} [\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t].$$

We refer to the value  $p_{t,T}^i$  as the *full-information value* of the security with a dividend stream  $D$ . The computation of the full information value  $p^i$  can be done in a number of ways; see the remark after Proposition 2.5.1.

On the other hand, the **market information**  $(\mathcal{F}_t^Z)_{t \geq 0}$  is given by  $\mathcal{F}_t^Z := \sigma\{Z_t : t \geq 0\}$ .<sup>2</sup> For an investor in the financial market, who knows only the observation process  $Z$  and has no access to the unobservable  $X$ , the market price of the dividend stream  $D$  is the expectation of the discounted payoff  $H_{t,T}$  conditional on the partial information  $\mathcal{F}_t^Z$  up to time  $t$ , i.e.,

$$\hat{p}_{t,T}^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] = \mathbb{E}^{\mathbb{Q}} [\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t^Z].$$

We call  $\hat{p}_{t,T}^i$  the *market price* of the security.

## 2.3 Practical Models

### 2.3.1 Model I: A Simple Model with One State Process

In this section, we present a simple example to illustrate the basic idea. We consider the case where there is only the global state process  $X = (X_t)_{t \geq 0}$ .

Suppose there exists a state process  $(X_t)_{t \geq 0}$  that drives the common dynamics of the credit derivatives of the  $N$  firms. More precisely,  $X := (X_t)_{t \geq 0}$  is a continuous-time, finite-state, hidden Markov process, defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with state space

---

<sup>2</sup>Note that we assumed  $(\mathcal{F}_t)_{t \geq 0}$  is rich enough, so that all stochastic processes are  $(\mathcal{F}_t)$ -adapted. In particular,  $\mathcal{F}_t := \sigma\{(X_s, Z_s) : s \leq t\}$ , and thus  $\mathcal{F}_t^Z \subseteq \mathcal{F}_t$ .

$S$ . Without loss of generality, we take the state space of  $X$  to be the set  $\{e_1, \dots, e_k\} \subseteq \mathbb{R}^k$  of unit vectors, where  $e_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^k$  with the “1” in the  $i$ th entry and where the superscript  $\top$  represents the transpose of a row vector. This is called the *canonical representation* of the state space of the Markov chain  $X$ .

Let  $A$  denote a constant rate matrix, or  $Q$ -matrix, of the Markov chain  $X$ . Then, using the canonical representation of the state space, the dynamics of the Markov chain  $X$  have the following semi-martingale representation

$$dX_t = AX_t dt + dM_t, \quad (2.2)$$

where  $(M_t)_{t \geq 0}$  is an  $\mathbb{R}^k$ -valued *martingale* with respect to the natural filtration generated by  $(X_t)_{t \geq 0}$ .

The default events of the  $N$  firms are assumed to depend solely on the process  $(X_t)_{t \geq 0}$ . The filtration  $\{\mathcal{F}_t, t \geq 0\}$  is defined by  $\mathcal{F}_t := \sigma\{X_s : s \leq t\}$ .

Recall that for  $1 \leq i \leq N$ , the random time  $\tau_i$  denotes the default time of firm  $i$ . Suppose for each firm  $i$ , the default intensity  $\lambda_t^i := \lambda^i(X_t)$  depends on the state process  $(X_t)_{t \geq 0}$ . The survival probability of the  $i$ th reference entity is given by

$$\mathbb{P}(\tau_i \geq t | \mathcal{F}_t) = \exp(-C_t^i),$$

where  $C_t^i$  is a functional of the form

$$C_t^i = \int_0^t \lambda^i(X_u) du.$$

Since the state space of  $X$  has the canonical representation, we can write  $\lambda^i = (\lambda_1^i, \lambda_2^i, \dots, \lambda_K^i)$ , so  $\lambda^i(X_t) = \langle \lambda^i, X_t \rangle$ .

Assume that conditional on the sample path of the chain, the defaults on the  $N$  firms are independent. Thus, for all  $t_1, \dots, t_N \geq 0$ ,

$$\mathbb{P}(\tau_1 > t_1, \dots, \tau_N > t_N | \mathcal{F}_t) = \prod_{i=1}^N \exp(-C_{t_i}^i).$$

Under this assumption, there are no joint defaults, i.e.,  $\tau_i \neq \tau_j$  for  $i \neq j$  almost surely. Also, almost surely  $X$  and  $Z$  do not jump together.

The loss state process  $Z$  is defined in the same way as in Section 2.1. The loss size  $\ell_i$  is a random variable with density function  $f^i(t, \ell)$ . Note that the density function  $f^i(t, \ell)$  also depends on the states  $X_{t-}$ , as is explained in Section 3.2.3.

For each  $k = 1, 2, \dots, K$ , write  $f_k^i(y)$  for a probability density function of the random jump size  $y^i = Z_u^i - Z_{u-}^i = \mathbf{1}_{\{u=\tau_i\}}\ell_t^i$ , when  $X_{u-} = e_k$ .

Suppose we have a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq K$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y)$ , independent of the hidden state  $X$ , for the loss sizes. The existence of such a reference probability measure  $\bar{\mathbb{P}}$  is shown in Elliott et al. (1995). (The main idea is that we can derive  $\mathbb{P}$  from  $\bar{\mathbb{P}}$ . Then working backward, we can obtain  $\bar{\mathbb{P}}$  from  $\mathbb{P}$ , and hence the existence of  $\bar{\mathbb{P}}$  is established).

For each  $k = 1, 2, \dots, K$ , define

$$h_k^i(y) = \frac{\lambda_k^i f_k^i(y)}{f(y)}.$$

Write  $H_k^i(u) := h_k^i(y_u)$ . Similarly to  $N_t$  defined in Section 3.2.3, we write  $N_t^i = \int_0^t \int_E \mu^i(du, de)$ .

**Theorem 2.3.1** *Let  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^K$ . Suppose  $q_t$  is a  $K$ -dimensional vector such that*

$$\begin{aligned} q_t = & q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \text{diag}(H_1^i(u) - 1, \dots, H_K^i(u) - 1) q_{u-} dN_u^i \\ & - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda_1^i - 1, \dots, \lambda_K^i - 1) q_u du. \end{aligned}$$

*Then*

$$\mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z] = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}.$$

*Proof.* See Appendix 1. □

**Remark.** With the above recursive SDE for  $q_t$ , we can use the Euler-Maruyama method, or the Euler-Milstein method, to obtain the approximate numerical solution for  $q$ . More

detailed discussion for methods for the approximate numerical solution of a SDE is given in Kloeden and Platen (1992). Once we have  $q$ , we can obtain an estimate of  $X$ .

With this theorem, we have an estimate of  $X$ , and then the fair price of a credit derivative can be obtained as follows.

### 2.3.2 Model II: A Practical Model with Contagion and Frailty

In section 3.2.1, we described a general ideal model. Ideally, the processes  $X, X^1, X^2, \dots, X^N$  will contain all underlying information relating to possible defaults of the firms. However, this model is complicated. To make the model more practical, we approximate the information contained in the Markov chains  $\{X^1, X^2, \dots, X^N\}$ .

In this section, we approximate  $\{X^1, X^2, \dots, X^N\}$  by a single continuous-time, finite-state Markov chain  $Y$ . In theory,  $\{X^1, X^2, \dots, X^N\}$  could be modeled with a single Markov chain  $Y$ . To see this, for  $i = 1, 2, \dots, N$ , let  $S_i$  be the state space of  $X^i$ , consisting of  $k_i$  states, i.e.,  $|S_i| = k_i$ . Then the state space of the process  $(Y_t)_{t \geq 0}$  would be

$$S = S_1 \times S_2 \times \dots \times S_N.$$

Thus, the process  $Y = (Y_t)_{t \geq 0}$  is a multi-dimensional vector process, each component containing information from  $X^i$ ,  $1 \leq i \leq N$ , and the state space  $S$  of  $(Y_t)_{t \geq 0}$  consists of  $K := k_1 k_2 \dots k_N$  states. Alternatively, each state of  $Y$  is a Cartesian product of the states of  $\{X^1, X^2, \dots, X^N\}$ . Therefore the process  $Y$  indeed contains all information of  $X^1, X^2, \dots, X^N$ , and the default intensity  $\lambda^i(X_t, X_t^1, X_t^2, \dots, X_t^N)$  in (2.1) can be given in terms of the states of  $X$  and  $Y$ .

Based on this idea, to make our model practical, we assume there are two independent, underlying Markov processes  $X$  and  $Y$ . Here  $X$  represents the global state, (e.g., the global economy, political regime, etc.), as in Section 3.2.1, and  $Y$  represents the firm-specific states for all firms  $i$ ,  $1 \leq i \leq N$ . The state process  $Y$  can be thought of as approximating the information contained in  $\{X^1, X^2, \dots, X^N\}$ .

Therefore, the processes  $Y$  models the “*contagion*”, and the process  $X$ , on the other hand, has the implications of “*frailty*”. Thus in this simple, practical model, both *contagion* and *frailty* are allowed.

The default events are assumed to depend on the processes  $(X_t)_{t \geq 0}$  and  $(Y_t)_{t \geq 0}$ . The default intensity  $\lambda_t^i := \lambda^i(X_t, Y_t)$  for firm  $i$  is now given by

$$\lambda^i(X_t, Y_t) = \alpha^i(X_t) + \beta^i(Y_t). \quad (2.3)$$

The function  $\beta^i(Y_t)$  can be thought to **approximate**  $\alpha_1^i(X_t^1) + \alpha_2^i(X_t^2) + \dots + \alpha_N^i(X_t^N)$  in (2.1).

Without loss of generality, we identify the state spaces of  $(X_t)_{t \geq 0}$  and  $(Y_t)_{t \geq 0}$  with the sets  $\{e_1, \dots, e_K\} \subseteq \mathbb{R}^K$  and  $\{\epsilon_1, \dots, \epsilon_M\} \subseteq \mathbb{R}^M$ , where  $e_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^K$  and  $\epsilon_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^M$  with the “1” in the  $i$ th entry. Then we have the following semi-martingale representations for  $X$  and  $Y$ :

$$\begin{aligned} dX_t &= A_t^1 X_t dt + dM_t^1 \in \mathbb{R}^K \\ dY_t &= A_t^2 Y_t dt + dM_t^2 \in \mathbb{R}^M, \end{aligned}$$

where  $(M_t^1)_{t \geq 0} \in \mathbb{R}^K$  and  $(M_t^2)_{t \geq 0} \in \mathbb{R}^M$  are two vector-valued martingales.

Since the states of  $X$  and  $Y$  are identified with unit vectors, we can write  $\alpha^i = (\alpha_1^i, \alpha_2^i, \dots, \alpha_K^i)$  and  $\beta^i = (\beta_1^i, \beta_2^i, \dots, \beta_M^i)$ , so that  $\alpha^i(X_t) = \langle \alpha^i, X_t \rangle$  and  $\beta^i(Y_t) = \langle \beta^i, Y_t \rangle$ . Then the intensity  $\lambda_t^i$  in equation (2.3) is given by

$$\lambda^i(X_t, Y_t) = \langle \alpha^i, X_t \rangle + \langle \beta^i, Y_t \rangle. \quad (2.3')$$

For each firm  $i$ ,  $1 \leq i \leq N$ , the default time  $\tau_i$  is assumed to be a random time, with

$$\mathbb{P}(\tau_i \geq t | \mathcal{F}_t) = \exp\left(-\int_0^t \lambda_s^i ds\right).$$

The loss state process  $Z$  is defined in the same way as in Section 2.1. The loss size  $\ell_i$  is a random variable with density function  $f^i(\tau_i, \ell)$ . Here the parameter  $\tau$  in the function  $f^i(\tau, \ell)$



indicates that the density may depend on the default time  $\tau$ . Note that the density function  $f^i(\tau, \ell)$  also depends on the states  $X_{\tau-}$  and  $Y_{\tau-}$ , as is explained in Section 3.2.3. For each  $k = 1, 2, \dots, K$ , and  $m = 1, 2, \dots, M$ , write  $f_{km}^i(y)$  for a probability density function of the random jump size  $y^i = \mathbf{1}_{\{u=\tau_i\}} \ell_u^i$ , when  $X_{u-} = e_k$  and  $Y_{u-} = \epsilon_m$ .

For  $1 \leq i \leq N$ , define  $Z_t^i := \ell_i \mathbf{1}_{\{\tau_i \leq t\}}$ . Then the *loss state* of the credit derivatives is given by the multi-dimensional vector process  $(Z_t)_{t \geq 0} = (Z_t^1, \dots, Z_t^N)$ .

Suppose we have a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq N$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y^i)$ , independent of the hidden states  $X_t$  and  $Y_t$ , for the loss sizes. For the existence of the reference probability measure  $\bar{\mathbb{P}}$ , the reader is referred to Elliott et al. (1995).

For each  $k = 1, 2, \dots, K$ ,  $m = 1, 2, \dots, M$  define

$$h_{km}^i(y^i) = \frac{(\alpha_k^i + \beta_m^i) f_{km}^i(y^i)}{f(y^i)}.$$

Write  $H_{km}^i(u) := h_{km}^i(y_u)$ .

In this section, we use  $\otimes$  to denote the *tensor product* (or *outer product*) of two vectors. The *outer product* of two vectors  $X$  and  $Y$  is defined to be  $X \otimes Y := XY^\top$ .

Note that by the definition of outer product,  $\mathbb{E}(X_t \otimes Y_t | \mathcal{F}_t^Z)$  is a  $K \times M$  matrix.

**Theorem 2.3.2** *Suppose  $q_t$  is a  $K \times M$  matrix such that*

$$\begin{aligned} q_t = q_0 &+ \int_0^t A^1 q_u du + \int_0^t q_u (A^2)^\top du + \sum_{i=1}^N \int_0^t \begin{pmatrix} H_{11}^i(u) - 1 & \dots & H_{1M}^i(u) - 1 \\ \vdots & \ddots & \vdots \\ H_{K1}^i(u) - 1 & \dots & H_{KM}^i(u) - 1 \end{pmatrix} \circ q_u - dN_u^i \\ &- \sum_{i=1}^N \int_0^{t \wedge \tau_i} \begin{pmatrix} (\alpha_1^i + \beta_1^i) - 1 & \dots & (\alpha_1^i + \beta_M^i) - 1 \\ \vdots & \ddots & \vdots \\ (\alpha_K^i + \beta_M^i) - 1 & \dots & (\alpha_K^i + \beta_M^i) - 1 \end{pmatrix} \circ q_u du, \end{aligned}$$

where the operation  $\circ$  denotes the *Hadamard product* (entrywise product) of two matrices.

Then

$$\mathbb{E}^{\mathbb{P}}[X_t \otimes Y_t | \mathcal{F}_t^Z] = \frac{q_t}{(\mathbf{1}_K)^\top q_t \mathbf{1}_M}.$$

*Proof.* See Appendix 2. □

With the above recursive SDE for  $q_t$ , we can use the Euler-Maruyama method, or the Euler-Milstein method, to obtain the approximate numerical solution for  $q$ , and then obtain the estimate for  $X \otimes Y$ .

### 2.3.3 Model III: A More Refined Model with Contagion and Frailty

In this section, we consider a more refined model than in Section 2.3.2. We approximate the information contained in  $\{X^1, X^2, \dots, X^N\}$  in Section 3.2.1 with two independent Markov chains  $Y^1$  and  $Y^2$ , rather than with one chain  $Y$  as in Section 2.3.2. In practice, approximating with two chains  $Y^1$  and  $Y^2$  may enable us to obtain better information than using one chain  $Y$  as in Section 2.3.2. Also, the Markov process  $X$  represents the global state (e.g., the global economy, political regime, etc.) or frailty as in Sections 3.2.1 and 2.3.2.

Without loss of generality, we identify the state spaces of  $(X_t)_{t \geq 0}$ ,  $(Y_t^1)_{t \geq 0}$  and  $(Y_t^2)_{t \geq 0}$  with the sets  $\{e_1, \dots, e_K\} \subseteq \mathbb{R}^K$ ,  $\{\epsilon_1, \dots, \epsilon_M\} \subseteq \mathbb{R}^M$  and  $\{\delta_1, \dots, \delta_J\} \subseteq \mathbb{R}^J$ , respectively, where  $e_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^K$ ,  $\epsilon_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^M$  and  $\delta_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^J$  with the “1” in the  $i$ th entry. Then we have the following semi-martingale representations for  $X$ ,  $Y^1$  and  $Y^2$ :

$$\begin{aligned} dX_t &= A_t^1 X_t dt + dM_t^1 \in \mathbb{R}^K \\ dY_t^1 &= A_t^2 Y_t^1 dt + dM_t^2 \in \mathbb{R}^M \\ dY_t^2 &= A_t^3 Y_t^2 dt + dM_t^3 \in \mathbb{R}^J, \end{aligned}$$

where  $(M_t^1)_{t \geq 0} \in \mathbb{R}^K$ ,  $(M_t^2)_{t \geq 0} \in \mathbb{R}^M$  and  $(M_t^3)_{t \geq 0} \in \mathbb{R}^J$  are three vector-valued martingales. The default events are assumed to depend on the processes  $(X_t)_{t \geq 0}$ ,  $(Y_t^1)_{t \geq 0}$  and  $(Y_t^2)_{t \geq 0}$ .

Suppose the default intensity  $\lambda_t^i := \lambda^i(X_t, Y_t^1, Y_t^2)$  is given by

$$\begin{aligned}\lambda^i(X_t, Y_t^1, Y_t^2) &= \alpha^i(X_t) + \beta^i(Y_t^1) + \zeta^i(Y_t^2) \\ &= \langle \alpha^i, X_t \rangle + \langle \beta^i, Y_t^1 \rangle + \langle \zeta^i, Y_t^2 \rangle.\end{aligned}$$

Again,  $\beta^i(Y_t^1) + \zeta^i(Y_t^2)$  can be considered an **approximation** of  $\alpha_1^i(X_t^1) + \alpha_2^i(X_t^2) + \dots + \alpha_N^i(X_t^N)$  in equation (2.1).

We can write  $\alpha^i = (\alpha_1^i, \alpha_2^i, \dots, \alpha_K^i)$ ,  $\beta^i = (\beta_1^i, \beta_2^i, \dots, \beta_M^i)$ , and  $\zeta^i = (\zeta_1^i, \zeta_2^i, \dots, \zeta_J^i)$ , such that  $\alpha^i(X_t) = \langle \alpha^i, X_t \rangle$ ,  $\beta^i(Y_t^1) = \langle \beta^i, Y_t^1 \rangle$ , and  $\zeta^i(Y_t^2) = \langle \zeta^i, Y_t^2 \rangle$ .

For each firm  $i$ ,  $1 \leq i \leq N$ , the default time  $\tau_i$  is assumed to be a random time, with

$$\mathbb{P}(\tau_i \geq t | \mathcal{F}_t) = \exp\left(-\int_0^t \lambda_s^i ds\right).$$

The loss state process  $Z$  is defined in the same way as in Section 2.1. The loss size  $\ell_i$  is a random variable with density function  $f^i(\tau, \ell)$ . Here the parameter  $t$  in the function  $f^i(\tau, \ell)$  indicates that the density may depend on the default time  $\tau$ . Note that the density function  $f^i(\tau, \ell)$  also depends on the states  $X_{\tau-}$ ,  $Y_{\tau-}^1$  and  $Y_{\tau-}^2$ , as is explained in Section 3.2.3. For each  $k = 1, 2, \dots, K$ ,  $m = 1, 2, \dots, M$  and  $j = 1, 2, \dots, J$ , write  $f_{kmj}^i(y)$  for a probability density function of the random jump size  $y^i = \mathbf{1}_{\{u=\tau_i\}} \ell_u^i$ , when  $X_{u-} = e_k$ ,  $Y_{u-}^1 = \epsilon_m$  and  $Y_{u-}^2 = \delta_j$ .

For  $1 \leq i \leq N$ , define  $Z_t^i := \ell_i \mathbf{1}_{\{\tau_i \leq t\}}$ . Then the *loss state* of the credit derivatives is given by the multi-dimensional vector process  $(Z_t)_{t \geq 0} = (Z_t^1, \dots, Z_t^N)$ .

From Elliott et al. (1995), we can suppose that there exists a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq N$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y^i)$ , independent of the hidden states  $X_t$  and  $Y_t^1, Y_t^2$ , for the loss sizes.

For each  $k = 1, 2, \dots, K$ ,  $m = 1, 2, \dots, M$  and  $j = 1, 2, \dots, J$ , define

$$h_{kmj}^i(y^i) = \frac{(\alpha_k^i + \beta_m^i + \zeta_j^i) f_{kmj}^i(y^i)}{f(y^i)}.$$

Write  $H_{kmj}^i(u) := h_{kmj}^i(y_u)$ .

In this section, let  $\odot$  denote the *Kronecker product* of two matrices. The Kronecker product has the following property. For matrices  $A, B, C, D$ ,

$$(A \odot B)(P \odot Q) = (AP) \odot (BQ).$$

Note that the notation  $\odot$  here denotes an operation different from  $\otimes$  in section 2.3.2

**Theorem 2.3.3** *Suppose  $q_t$  is a  $(KMJ)$ -dimensional vector such that*

$$\begin{aligned} q_t = & q_0 + \int_0^t (A^1 \odot I_M \odot I_J) q_u du + \int_0^t (I_K \odot A^2 \odot I_J) q_u du + \int_0^t (I_K \odot I_M \odot A^3) q_u du \\ & + \sum_{i=1}^N \int_0^t (H_{111}^i(u) - 1, \dots, H_{11J}^i(u) - 1, \dots, H_{KMJ}^i(u) - 1)^\top \circ q_{u-} \cdot dN_u^i \\ & - \sum_{i=1}^N \int_0^{t \wedge \tau_i} ((\alpha_1^i + \beta_1^i + \zeta_1^i) - 1, \dots, (\alpha_1^i + \beta_1^i + \zeta_J^i) - 1, \dots, (\alpha_K^i + \beta_M^i + \zeta_J^i) - 1)^\top \circ q_u du, \end{aligned}$$

where the operation  $\circ$  denotes the *Hadamard product* (entrywise product) of two matrices.

Then

$$\mathbb{E}^\mathbb{P}[X_t \odot Y_t^1 \odot Y_t^2 | \mathcal{F}_t^Z] = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}.$$

*Proof.* See Appendix 3. □

With the above recursive SDE for  $q_t$ , we can use the Euler-Maruyama method, or the Euler-Milstein method, to obtain the approximate numerical solution for  $q$ , and then obtain the estimate of  $X \odot Y^1 \odot Y^2$ .

## 2.4 Default Probabilities

We wish to find the conditional joint default probability of the firms given observed information  $\mathcal{F}_t^Z$  up to time  $t$ , and given that the firms have not defaulted yet by time  $t$ .

With the filtering result given in Theorem 2.3.1, we are able to obtain the following closed-form expression for the default probability of the firms given  $\mathcal{F}_t^Z$ , and given that the firms have not defaulted yet by time  $t$ .

**Theorem 2.4.1** For each  $i$ ,  $1 \leq i \leq N$ , and  $t, t+h \in [0, T]$  with  $h > 0$ , given  $\mathcal{F}_t^Z$  and that the firms have not defaulted by time  $t$ .

$$\begin{aligned} & \mathbb{P}(t < \tau_i < t+h | \mathcal{F}_t^Z) \\ &= \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} \int_{[t, t+h]^K} \left( 1 - e^{-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) \\ & \quad e^{-\sqrt{-1} \langle \xi, \mathbf{O}_{t,h} \rangle} \langle \exp((A + \sqrt{-1} \text{diag}(\xi)h) \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z], \mathbf{1}) d(\xi). \end{aligned}$$

*Proof.* Let  $O_{t,h}^i$  be the occupation time of the chain  $X$  in state  $e_i$  over the time interval  $[t, t+h]$ , that is,

$$O_{t,h}^i := \int_t^{t+h} \langle X_u, e_i \rangle du.$$

Write

$$\mathbf{O}_{t,h} := (O_{t,h}^1, O_{t,h}^2, \dots, O_{t,h}^K) \in [t, t+h]^K.$$

For any  $i$ ,  $1 \leq i \leq N$ ,

$$\begin{aligned} \mathbb{P}(t < \tau_i < t+h | \mathcal{F}_t^Z) &= \mathbb{E}^{\mathbb{P}} \left[ \left( 1 - \exp \left( - \int_t^{t+h} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \right) \right) \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{P}} \left[ (1 - \exp(-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle) | \mathcal{F}_t) \right] \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \int_{[t, t+h]^K} \left( 1 - \exp(-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle) \right) \phi(\mathbf{O}_{t,h}) d(\mathbf{O}_{t,h}) \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \int_{[t, t+h]^K} \left( 1 - e^{-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) \phi(\mathbf{O}_{t,h}) d(\mathbf{O}_{t,h}) \middle| \mathcal{F}_t^Z \right], \end{aligned}$$

where  $\phi(\mathbf{O}_{t,h})$  is the conditional joint density function of  $\mathbf{O}_{t,h}$  given  $\mathcal{F}_t$  under  $\mathbb{P}$ . Since the conditional joint characteristic function of  $\mathbf{O}_{t,h}$  given  $\mathcal{F}_t$ , evaluated at  $\xi \in \mathbb{R}^K$  under  $\mathbb{P}$  is

$$\begin{aligned} \Phi_{\mathbf{O}_{t,h}}(\xi) &:= \mathbb{E}^{\mathbb{P}} \left[ e^{\sqrt{-1} \langle \xi, \mathbf{O}_{t,h} \rangle} | \mathcal{F}_t \right] \\ &= \exp((A + \sqrt{-1} \text{diag}(\xi)h) X_t, \mathbf{1}), \end{aligned}$$

then the conditional joint density function  $\phi(\mathbf{O}_{t,h})$  of  $\mathbf{O}_{t,h}$  given  $\mathcal{F}_t$  under  $\mathbb{P}$  is given by the inverse Fourier transform

$$\phi(\mathbf{O}_{t,h}) = \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} e^{-\sqrt{-1} \langle \xi, \mathbf{O}_{t,h} \rangle} \Phi_{\mathbf{O}}(\xi) d(\xi)$$

$$= \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} e^{-\sqrt{-1}\langle \xi, \mathbf{O}_{t,h} \rangle} \langle \exp((A + \sqrt{-1} \text{diag}(\xi)h) X_t, \mathbf{1}) d(\xi).$$

Thus we have

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}} \left[ \left( 1 - \exp \left( - \int_t^{t+h} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \right) \right) \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \int_{[t, t+h]^K} \left( 1 - e^{-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) \phi(\mathbf{O}_{t,h}) d(\mathbf{O}_{t,h}) \middle| \mathcal{F}_t^Z \right] \\ &= \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} \int_{[t, t+h]^K} \left( 1 - e^{-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) e^{-\sqrt{-1}\langle \xi, \mathbf{O}_{t,h} \rangle} \langle \exp((A + \sqrt{-1} \text{diag}(\xi)h) \widehat{X}_t, \mathbf{1}) d(\mathbf{O}_{t,h}) d(\xi) \\ &= \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} \int_{[t, t+h]^K} \left( 1 - e^{-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) \\ & \quad e^{-\sqrt{-1}\langle \xi, \mathbf{O}_{t,h} \rangle} \langle \exp((A + \sqrt{-1} \text{diag}(\xi)h) \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z], \mathbf{1}) d(\mathbf{O}_{t,h}) d(\xi). \end{aligned}$$

Then

$$\begin{aligned} & \mathbb{P}(t < \tau_i < t+h | \mathcal{F}_t^Z) \\ &= \mathbb{E}^{\mathbb{P}} \left[ \left( 1 - \exp \left( - \int_t^{t+h} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \right) \right) \middle| \mathcal{F}_t^Z \right] \\ &= \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} \int_{[t, t+h]^K} \left( 1 - e^{-\langle \lambda^i(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) \\ & \quad e^{-\sqrt{-1}\langle \xi, \mathbf{O}_{t,h} \rangle} \langle \exp((A + \sqrt{-1} \text{diag}(\xi)h) \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z], \mathbf{1}) d(\mathbf{O}_{t,h}) d(\xi). \end{aligned}$$

The desired formula is proven.  $\square$

Similarly, we can get the joint default probability for any combination of  $m$  firms  $i_1, i_2, \dots, i_m$ .

**Theorem 2.4.2** For  $t, t+h \in [0, T]$  with  $h > 0$ , given  $\mathcal{F}_t^Z$  and that the firms  $i_1, i_2, \dots, i_m$  have not defaulted by time  $t$ ,

$$\begin{aligned} & P(t \leq \tau_{i_1} \leq t+h, t \leq \tau_{i_2} \leq t+h, \dots, t \leq \tau_{i_m} \leq t+h | \mathcal{F}_t^Z) \\ &= \frac{1}{(2\pi)^K} \int_{\mathbb{R}^K} \int_{[t, t+h]^K} \prod_{j=1}^m \left( 1 - e^{-\langle \lambda^{i_j}(u, X_{u-}, N_{(\cdot)}), \mathbf{O}_{t,h} \rangle} \right) \\ & \quad \times \langle \exp((A + \sqrt{-1} \text{diag}(\xi)h) \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z], \mathbf{1}) e^{-\sqrt{-1}\langle \mathbf{o}, \xi \rangle} d(O_{t,h}^1, O_{t,h}^2, \dots, O_{t,h}^K) d(\xi) \end{aligned}$$

*Proof.* Similar to that of Theorem 2.4.1.  $\square$

## 2.5 Pricing of Contingent Claims

The expected default probabilities obtained in Section 2.4 cannot be used directly to the pricing of credit-risk bearing securities.<sup>3</sup> These quantities are in the historical measure  $\mathbb{P}$ , while the pricing of the securities needs to be carried out in a risk-neutral measure  $\mathbb{Q}$ .

In order to obtain the price of credit-risk bearing securities, we need to define a risk-neutral measure. We follow the principle presented in Duffie and Singleton (2003), and Dai and Singleton (2003) to consider the pricing. For the pricing and related parameter estimation, as customary, three key ingredients are needed:

- $\mathbb{P}$ : the historical (physical, or “real world”) measure, as well as the time-series and parameters in this measure;
- $\mathbb{Q}$ : the risk-neutral (or the pricing) measure, as well as the time-series and parameters in this measure; we shall implement the pricing of the securities in this measure.
- $M_t$ : the pricing kernel (stochastic discount factor) that links the historical measure  $\mathbb{P}$  and risk-neutral measure  $\mathbb{Q}$ . Usually  $M_t = \exp(-\int_0^t r_u du)\Lambda_t$  where  $\Lambda_t$  is the the Radon-Nikodym derivative  $d\mathbb{Q}/d\mathbb{P}$ .

As is stressed repeatedly in Duffie and Singleton (2003), Dai and Singleton (2003), the pricing of the securities will be carried out in the risk-neutral measure  $\mathbb{Q}$ , and the parameter estimation needs to be done in the historical measure  $\mathbb{P}$ .

---

<sup>3</sup>The expected default probabilities and expected recovery rates are obtained from historical frequencies of defaults and rates of recoveries and, therefore, these magnitudes are under the historical measure  $\mathbb{P}$ . If one uses these expected probability and recoveries under the historical measure  $\mathbb{P}$  to estimate the price of the securities, the implied yield spread is called an *actuarial credit spread*. Usually there exist big differences between the *actuarial* credit spreads and the actual (market) yield spreads. The differences are due to many effects, including default risk premium, tax shields, and liquidity effects, etc. (see, e.g., Elton et al. (2001), Duffie and Singleton (2003)).

### 2.5.1 Price of a Contingent Claim in Model I

Consider a portfolio of  $N$  credit derivatives with common maturity  $T$ , with *cumulative* dividend streams  $D_t^i$  given.  $1 \leq i \leq N$ . The discounted cumulative dividend streams are

$$\tilde{D}_t^i := \int_0^t e^{-\int_0^s r_u du} dD_s^i.$$

Then the discounted payoff of a credit derivative when it matures is (it is indeed the total sum of discounted payments in the time interval  $(t, T]$ )

$$H_{t,T}^i := \int_t^T e^{-\int_0^s r_u du} dD_s = \int_t^T d\tilde{D}_s = \tilde{D}_T^i - \tilde{D}_t^i.$$

We can imagine that for an econometrician who knows both the observation  $Z$  and the unobservable latent factor  $X$ , the value of a security with dividend stream  $D$  is the expectation of the discounted payoff  $H_{t,T}$  conditional on the full information  $\mathcal{F}_t$  up to time  $t$ , i.e.,

$$p_{t,T}^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} [\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t].$$

We refer to the value  $p_{t,T}^i$  as the *full-information value* of the security with a dividend stream  $D$ .

On the other hand, for an investor who knows only the observation process  $Z$  and has no access to the unobservable  $X$ , the market price of the dividend stream  $D$  is the expectation of the discounted payoff  $H_{t,T}$  conditional on the partial information  $\mathcal{F}_t^Z$  up to time  $t$ , i.e.,

$$\hat{p}_{t,T}^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] = \mathbb{E}^{\mathbb{Q}} [\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t^Z].$$

We call  $\hat{p}_{t,T}^i$  the *market price* of the security. Note that  $\hat{p}_{t,T}^i$  depends on the current state  $X_t$  and  $Z_t$ .<sup>4</sup> In the following theorem, we derive the expression for the market price of a security with a dividend stream  $D$ .

---

<sup>4</sup>From the Markov property of the processes  $(X, Z)$ , we have (see, e.g., Chapter 2, Section 5 in Karatzas and Shreve (1988))  $\mathbb{E}^{\mathbb{Q}} [\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t^Z] = p^i(t, X_t, Z_t)$ , i.e.,  $p_t^i$  can be written as a function of  $X_t$  and  $Z_t$ .



**Proposition 2.5.1** *The market price  $\widehat{p}_{t,T}^i$  of a credit derivative with dividend stream  $D_t^i$  is*

$$\begin{aligned}\widehat{p}_{t,T}^i &= \mathbb{E}^{\mathbb{Q}} [p^i(t, X_t, Z_t) | \mathcal{F}_t^Z] \\ &= \left( p_{t,T}^i(e_1, Z_t), p_{t,T}^i(e_2, Z_t), \dots, p_{t,T}^i(e_K, Z_t) \right) \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z].\end{aligned}$$

*Proof.* Note that  $\mathbb{E}(X_t | \mathcal{F}_t^Z)$  is a  $K$ -dimensional vector, which is given in Theorem 2.3.1. By the definition of  $\widehat{p}_t^i$ , we have

$$\begin{aligned}\widehat{p}_t^i &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \Lambda_t \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{P}} \left[ \Lambda_t \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \mathbb{E}^{\mathbb{P}} [X | \mathcal{F}_t^Z] \\ &= \left( p_t^i(e_1, Z_t), p_t^i(e_2, Z_t), \dots, p_t^i(e_K, Z_t) \right) \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z].\end{aligned}$$

Here, in the last equation, note that the product is the inner product of two  $K$ -dimensional vectors. □

**Remark.** The computation of the full-information value  $p^i$  has been studied by several researchers, and can be done in a number of ways. For example, it has been investigated in Elliott and Mamon (2003), using Markov-chain techniques. It is also considered in Graziano and Rogers (2009), using Laplace transforms. Thus, once the parameters of the model are determined, we can use these techniques to find the full-information value  $p_t^i$ . Then the market price  $\widehat{p}_t^i$  is obtained by an inner product of  $p_t^i$  and  $\mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z]$ , which is given in Theorem 2.3.1.

We now give the pricing formulars for some simple examples. For a detailed discussion of various contingent claims, see Duffie and Singleton (2003), and Lando (2004). With

Proposition 2.5.1, and arguments similar to those in Lando (2004), we can derive the following pricing formulas.

- First, consider a claim paying  $g(X_s)\mathbf{1}_{\{\tau>s\}}$  continuously until default or maturity time  $T$  when no default happens prior to  $T$ . Then for this claim, the market price is

$$\hat{p}_{t,T}^i = \mathbf{1}_{\{\tau>t\}} \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T g(X_s) e^{-\int_t^s (r_u + \lambda(X_u)) du} ds \middle| \mathcal{F}_t \right] \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z].$$

- Another typical example is a claim paying  $h(X_\tau)$  at default time  $\tau$ . The market price of this claim is

$$\hat{p}_{t,T}^i = \mathbf{1}_{\{\tau>t\}} \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T h(X_s) \lambda(X_s) e^{-\int_t^s (r_u + \lambda(X_u)) du} ds \right] \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z].$$

- A defaultable corporate bond can be decomposed as a sum of the above two types of claims. In particular, the price of a zero-coupon defaultable bond is (assuming zero-recovery)

$$\hat{p}_{t,T}^i = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_t^s (r_u + \lambda(X_u)) du} ds \middle| \mathcal{F}_t \right] \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z].$$

- When a zero-coupon defaultable bond has a fractional recovery rate  $\delta = \ell^i$  of market value when default, the price is

$$\hat{p}_{t,T}^i = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_t^s (r_u + (1-\delta)\lambda(X_u)) du} ds \middle| \mathcal{F}_t \right] \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z].$$

## 2.5.2 Price of a Contingent Claim in Model II

Since we now have the estimate of  $X$  and  $Y$ , we can obtain the fair price of a credit derivative in the following proposition. We note that  $\tilde{D}_T^i - \tilde{D}_t^i$  depends on the path  $\{Z_u : t < u \leq T\}$ . Also, by definition, the hypothetical value  $p_t^i$  depends on the states  $X_t$  and  $Y_t$ . Hence, similarly to Section 2.3.1, by the Markov property of the processes  $(X, Y, Z)$   $p_t^i$  can be written as a function of  $X_t$ ,  $Y_t$  and  $Z_t$ , i.e.,  $p_t^i = p_t^i(X_t, Y_t, Z_t)$ . For the computation of the full information value  $p^i$ , see the remark after Proposition 2.5.1.

**Proposition 2.5.2** *The market value  $\widehat{p}_t^i$  of the credit derivatives can be calculated by*

$$\widehat{p}_t^i = (1, \dots, 1) \begin{pmatrix} p_t^i(e_1, \epsilon_1, Z_t) & p_t^i(e_1, \epsilon_2, Z_t) & \dots & p_t^i(e_1, \epsilon_M, Z_t) \\ \vdots & \vdots & \ddots & \vdots \\ p_t^i(e_K, \epsilon_1, Z_t) & p_t^i(e_K, \epsilon_2, Z_t) & \dots & p_t^i(e_K, \epsilon_M, Z_t) \end{pmatrix} \circ \mathbb{E}^{\mathbb{P}}(X_t \otimes Y_t | \mathcal{F}_t^Z) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

where the operation  $\circ$  denotes the Hadamard product (entrywise product) of two matrices.

*Proof.* By the definition of  $\widehat{p}_t^i$ , we have

$$\begin{aligned} \widehat{p}_t^i &= \mathbb{E}^{\mathbb{Q}} [p_t^i(X_t, Y_t, Z_t) | \mathcal{F}_t^Z] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \Lambda_t \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{P}} \left[ \Lambda_t \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] \\ &= (1, \dots, 1) \begin{pmatrix} p_t^i(e_1, \epsilon_1, Z_t) & p_t^i(e_1, \epsilon_2, Z_t) & \dots & p_t^i(e_1, \epsilon_M, Z_t) \\ \vdots & \vdots & \ddots & \vdots \\ p_t^i(e_K, \epsilon_1, Z_t) & p_t^i(e_K, \epsilon_2, Z_t) & \dots & p_t^i(e_K, \epsilon_M, Z_t) \end{pmatrix} \circ \mathbb{E}^{\mathbb{P}}[X_t \otimes Y_t | \mathcal{F}_t^Z] \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \end{aligned}$$

where the operation  $\circ$  denotes the Hadamard product (entrywise product) of two matrices.

Here, in the last equation, note that the product of all the matrices and vectors gives a scalar.  $\square$

### 2.5.3 Price of a Contingent Claim in Model III

Now we have the estimate of  $X$ ,  $Y_1$  and  $Y_2$ . From this estimate, we can obtain the fair price of a credit derivative in the following proposition. We note that the hypothetical value  $p_t^i$  depends on the states  $X_t$ ,  $Y_t^1$  and  $Y_t^2$ . As in Section 2.3.1,  $p_t^i$  can be written as a function of  $X_t$ ,  $Y_t^1, Y_t^2$  and  $Z_t$ , i.e.,  $p_t^i = p_t^i(X_t, Y_t^1, Y_t^2, Z_t)$ . For the computation of the full information value  $p^i$ , see the remark after Proposition 2.5.1.

**Proposition 2.5.3** *The market value  $\widehat{p}_t^i$  of a credit derivative with dividend stream  $D_t^i$  is*

$$\begin{aligned}\widehat{p}_t^i &= \mathbb{E}^{\mathbb{Q}}(p_t^i(X_t, Y_t^1, Y_t^2, Z_t) | \mathcal{F}_t^Z) \\ &= (p_t^i(e_1, \epsilon_1, \delta_1, Z_t), \dots, p_t^i(e_1, \epsilon_1, \delta_J, Z_t), \dots, p_t^i(e_1, \epsilon_M, \delta_J, Z_t), \dots, p_t^i(e_K, \epsilon_M, \delta_J, Z_t)) \\ &\quad \cdot \mathbb{E}^{\mathbb{P}}(X_t \odot Y_t^1 \odot Y_t^2 | \mathcal{F}_t^Z).\end{aligned}$$

*Proof.* Note that by the definition of Kronecker product,  $\mathbb{E}^{\mathbb{Q}}[X_t \otimes Y_t^1 \otimes Y_t^2 | \mathcal{F}_t^Z]$  is a  $(KMJ)$ -dimensional vector. By the definition of  $\widehat{p}_t^i$ , we have

$$\begin{aligned}\widehat{p}_t^i &= \mathbb{E}^{\mathbb{Q}}(D_T^i - D_t^i | \mathcal{F}_t^Z) \\ &= \mathbb{E}^{\mathbb{P}}(\Lambda_t(D_T^i - D_t^i | \mathcal{F}_t) | \mathcal{F}_t^Z) \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \mathbb{E}^{\mathbb{P}} [X | \mathcal{F}_t^Z] \\ &= (p_t^i(e_1, \epsilon_1, \delta_1, Z_t), \dots, p_t^i(e_1, \epsilon_1, \delta_J, Z_t), \dots, p_t^i(e_1, \epsilon_M, \delta_J, Z_t), \dots, p_t^i(e_K, \epsilon_M, \delta_J, Z_t)) \\ &\quad \cdot \mathbb{E}^{\mathbb{P}}(X_t \odot Y_t^1 \odot Y_t^2 | \mathcal{F}_t^Z).\end{aligned}$$

Here, in the last equation, note that the product is the inner product of two  $(KMJ)$ -dimensional vectors. □

## 2.6 Dynamic Hedging

In this section, we investigate the dynamic hedging strategies for the credit derivatives in our model.

### 2.6.1 Scheme of Hedging

We follow the form for hedging strategies considered in Föllmer and Sondermann (1986), Föllmer and Schweizer (1990), Elliott and Föllmer (1991). In this section, we summarize the framework.

For the purpose of demonstration, in the following description we assume there is only one risky asset. All results still hold true when there are  $N > 1$  risky assets, i.e., when  $\theta_t$  is a multi-dimensional vector rather than a scalar.

Let  $(\Omega, \mathcal{F}, \mathbb{Q})$  be a probability space, and  $\{\mathcal{F}_t\}_{t \geq 0}$  a filtration. The probability measure  $\mathbb{Q}$  is assumed to be the risk neutral measure. Also, for simplicity, we work directly with discounted prices of the securities. Therefore, the discounted money market (default-free bond) account  $\{\tilde{B}_t\}_{t \geq 0}$  is such that  $\tilde{B}_t \equiv 1$ , for all  $t \geq 0$ , and hence  $d\tilde{B}_t \equiv 0$ . Suppose the martingale  $\tilde{X}_t$  represents the discounted price of some primitive (underlying) asset at time  $t$ .

Consider a trading strategy  $(\theta, \eta) := \{(\theta_t, \eta_t)\}_{t \geq 0}$ , which means that at time  $t$  the investor invests amount  $\theta_t$  in the primitive asset, and an amount  $\eta_t$  in the riskless money market account (default-free bond) with zero interest rate. For a trading strategy  $\{(\theta_t, \eta_t)\}_{t \geq 0}$ , the discounted *value* of the portfolio at time  $t$  is

$$\tilde{V}_t = \theta_t \tilde{X}_t + \eta_t \tilde{B}_t.$$

The strategy  $\{(\theta_t, \eta_t)\}_{t \geq 0}$  is said to be *self-financing*, if for all  $t \geq 0$ ,

$$d\tilde{V}_t = \theta_t d\tilde{X}_t + \eta_t d\tilde{B}_t.$$

Therefore, a trading strategy  $(\theta, \eta)$  gives rise to the following processes:

- a (discounted) value process  $\tilde{V}_t$ , which is defined to be

$$\begin{aligned} \tilde{V}_t &= \theta_t \tilde{X}_t + \eta_t \tilde{B}_t \\ &= \theta_t \tilde{X}_t + \eta_t, \end{aligned} \quad (\tilde{B}_t \equiv 1)$$

- a (discounted) gains process  $\tilde{G}_t$ , which is defined by

$$\begin{aligned} \tilde{G}_t &= \int_0^t \theta_s d\tilde{X}_s + \int_0^t \eta_s d\tilde{B}_s \\ &= \int_0^t \theta_s d\tilde{X}_s, \end{aligned} \quad (d\tilde{B}_t \equiv 0)$$

- a cumulative (discounted) cost process  $\tilde{C}_t$ , which is given by

$$\begin{aligned}\tilde{C}_t &= \tilde{V}_t - \tilde{G}_t \\ &= \tilde{V}_t - \int_0^t \theta_s d\tilde{X}_s.\end{aligned}$$

We note that the trading strategy  $\{(\theta_t, \eta_t)\}_{t \geq 0}$  is self-financing if, for all  $t \geq 0$ ,  $d\tilde{V}_t = \theta_t d\tilde{X}_t + \eta_t d\tilde{B}_t = d\tilde{G}_t$ . This is equivalent to  $\tilde{C}_t \equiv \tilde{C}_0 = V_0$ , for all  $t \geq 0$ .

If the security market is complete, then any contingent claim can be replicated by self-financing strategies. In many settings, the security market is incomplete, in the sense that not all contingent claims are attainable with self-financing strategies. In Föllmer and Sondermann (1986), Föllmer and Schweizer (1990), Föllmer et al. provide a way of measuring risk when the market is incomplete.

A trading strategy  $(\theta, \eta) := \{(\theta_t, \eta_t)\}_{t \geq 0}$  is said to be *mean-self-financing* if the corresponding discounted cost process  $\tilde{C} = \{\tilde{C}_t\}_{t \geq 0}$  is a martingale. In Föllmer et al. (1986) p. 209, the authors show that a trading strategy is mean-self-financing if and only if the corresponding discounted value process  $V = \{\tilde{V}_t\}_{t \geq 0}$  is a martingale.

Recall that a trading strategy  $(\theta, \eta)$  is self-financing if the corresponding discounted cost process  $\tilde{C} = \{\tilde{C}_t\}_{t \geq 0}$  has constant sample paths, i.e.,  $\tilde{C}_t \equiv \tilde{C}_0$ ,  $t \geq 0$ . Clearly, by definition, a self-financing strategy is also mean-self-financing, but not vice versa.

For any trading strategy  $(\theta, \eta)$ , denote the cumulative discounted cost process by  $\tilde{C} = \{\tilde{C}_t\}_{t \geq 0}$ . Then the *remaining risk* at time  $t$  is

$$\mathbb{E}^{\mathbb{Q}}[(\tilde{C}_T(\theta) - \tilde{C}_t(\theta))^2 | \mathcal{F}_t].$$

The remaining risk measures the intrinsic risk of a hedging strategy for a contingent claim at any time  $t$ .

Suppose that for  $T > 0$ , a contingent claim is given by  $\tilde{H}(X_T)$ , where  $\tilde{H}(\cdot)$  is a function such that  $\tilde{H}(X_T)$  is a real, square integrable random variable.

In Elliott and Föllmer (1991), the authors relate the hedging of the discounted contingent claim  $\tilde{H}(X_T)$  to the martingale representation of  $\tilde{H}(X_T)$ , and provide the following results (see Elliott and Föllmer (1991), pp. 141–143):

1. There exists a self-financing strategy  $(\theta, \eta)$  to replicate  $\tilde{H}(X_T)$  if and only if  $\tilde{H}(X_T)$  has the representation

$$\tilde{H}(X_T) = \mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)] + \int_0^T \xi_s d\tilde{X}_s, \quad \text{a.s.}$$

for some predictable integrand  $\xi$ , which is in turn equivalent to

$$\mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)] + \int_0^t \xi_s d\tilde{X}_s, \quad \text{a.s.}$$

2. There exists a *mean-self-financing strategy*  $(\theta, \eta)$  to replicate  $\tilde{H}(X_T)$  if and only if  $\tilde{H}(X_T)$  has the representation

$$\tilde{H}(X_T) = \mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)] + K_T + \int_0^T \xi_s d\tilde{X}_s, \quad \text{a.s.}$$

for some predictable integrand  $\xi$ . This is in turn equivalent to

$$\mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)] + K_t + \int_0^t \xi_s d\tilde{X}_s, \quad \text{a.s.}$$

where  $\{K_t\}_{t \geq 0}$  is a martingale.

3. There exists a *mean-self-financing strategy*  $(\theta, \eta)$  replicating  $\tilde{H}(X_T)$ , such that the *remaining risk*  $\mathbb{E}^{\mathbb{Q}}[(\tilde{C}_T(\theta) - \tilde{C}_t(\theta))^2|\mathcal{F}_t]$  is minimized, if and only if  $\mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)|\mathcal{F}_t]$  has the Kunita-Watanabe orthogonal representation (decomposition)

$$\mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}[\tilde{H}(X_T)] + \Gamma_t + \int_0^t \xi_s d\tilde{X}_s, \quad \text{a.s.}$$

for some predictable integrand  $\xi$ , where  $\Gamma = \{\Gamma_t\}_{t \geq 0}$  is a martingale orthogonal to  $\tilde{X}$ .

## 2.6.2 Hedging Strategies for the Credit Derivatives

As described in Bielecki and Rutkowski (2002), and Section 3.2.4, basically all credit derivatives in practice fall into two classes:

- Options on the firms' loss states. This class of derivatives has the cumulative dividend stream  $D_t$ , which can be written as (see Section 3.2.4)

$$\begin{aligned} D_t &= (Y\mathbf{1}_{\{\tau>T\}} + Y'\mathbf{1}_{\{\tau\leq T\}})\mathbf{1}_{\{t\geq T\}} + \int_0^t (1 - N_u)dC_u + \int_0^t R_u dN_u \\ &= (Y\mathbf{1}_{\{\tau>T\}} + Y'\mathbf{1}_{\{\tau\leq T\}})\mathbf{1}_{\{t\geq T\}} + \int_0^{t\wedge\tau} dC_u + R_\tau\mathbf{1}_{\{\tau\leq t\}}. \end{aligned}$$

Such derivatives include, e.g., corporate bonds, CDS, portfolio derivatives or basket derivatives, like collateralized debt obligations (CDO tranche), basket swaps, etc.

- Options on other derivatives. This class contains credit derivatives for which the payoff depends on the future market value of some other derivatives. Such derivatives include, e.g., options on corporate bonds, options on CDS indices, or options on portfolio CDO tranches, options on basket swaps, etc.

For the credit derivatives, we carry out our hedging strategies as follows. We consider primitive, underlying assets with prices  $\hat{\mathbf{p}} := \{\hat{\mathbf{p}}_t\}_{t\geq 0} = \{(p_t^1, \dots, p_t^N)\}_{t\geq 0}$ , which are derived in Section 3.2.5, and in Propositions 2.5.1, 2.5.2, 2.5.3. In general, as defined in Section 3.2.5, the ex-dividend price is,

$$\hat{p}_t^i := \mathbb{E}^{\mathbb{Q}}(\tilde{H}_{t,T}^i | \mathcal{F}_t^Z) = \mathbb{E}^{\mathbb{Q}}(\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t^Z).$$

If the quantity  $\tilde{H}_{t,T}^i(\omega)$  were to depend only on  $T$  and the sample path  $\omega \in \Omega$ , then  $\hat{p}_t^i := \mathbb{E}^{\mathbb{Q}}(\tilde{H}_{t,T}^i | \mathcal{F}_t^Z)$  would be a martingale. Then we can directly use the hedging scheme in Elliott and Föllmer (1991), as summarized in Section 2.6.1.

By definition, however,  $\tilde{H}_{t,T}^i(\omega)$  also depends on  $t$ , and hence  $\mathbb{E}^{\mathbb{Q}}(\tilde{H}_{t,T}^i | \mathcal{F}_t^Z) = \mathbb{E}^{\mathbb{Q}}(\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t^Z)$  is not necessarily a martingale. Thus the hedging scheme in Section 2.6.1 (or see



Föllmer and Sondermann (1986), Föllmer and Schweizer (1990), Elliott and Föllmer (1991) for more details) cannot be applied directly.

To get around this difficulty, we consider the *cumulative dividend prices* (*cum-dividend prices*) of the securities

$$\begin{aligned}\widehat{S}_t^i &:= \widehat{p}_t^i + \widetilde{D}_t^i = \mathbb{E}^{\mathbb{Q}}(\widetilde{H}_{t,T}^i | \mathcal{F}_t^Z) + \widetilde{D}_t^i \\ &= \mathbb{E}^{\mathbb{Q}}(\widetilde{D}_T^i - \widetilde{D}_t^i | \mathcal{F}_t^Z) + \widetilde{D}_t^i \\ &= \mathbb{E}^{\mathbb{Q}}(\widetilde{D}_T^i | \mathcal{F}_t^Z).\end{aligned}$$

Since the quantity  $\widetilde{D}_T^i(\omega)$  depends only on  $T$  and the sample path  $\omega \in \Omega$ , then with the local square integrability assumption,  $\widehat{S}_t^i := \mathbb{E}^{\mathbb{Q}}(\widetilde{D}_T^i | \mathcal{F}_t^Z)$  is a martingale. We use the hedging scheme in Elliott and Föllmer (1991), as summarized in Section 2.6.1, to the cum-dividend price  $\widehat{S}_t^i$ . Define  $\widehat{\mathbf{S}} := \{\widehat{\mathbf{S}}_t\}_{t \geq 0} = \{(\widehat{S}_t^1, \dots, \widehat{S}_t^N)\}_{t \geq 0}$ .

Suppose there is a claim  $H$  with a cumulative discounted dividend stream  $\widetilde{D} = (\{\widetilde{D}_t\}_{t \geq 0})$ . The claim  $H$  is the credit derivative we would like to hedge. The cum-dividend price process of  $H$  is  $\widehat{S}_t = \widehat{p}_t + \widetilde{D}_t = \mathbb{E}^{\mathbb{Q}}(\widetilde{D}_T | \mathcal{F}_t^Z)$ , a martingale.

Consider a trading strategy  $(\theta, \eta) := \{(\theta_t, \eta_t)\}_{t \geq 0} = \{(\theta_t^1, \dots, \theta_t^N, \eta_t)\}_{t \geq 0}$ , which means that at time  $t$  the investor invests amount  $\theta_t^i$  in the  $i$ th primitive, underlying asset, and an amount  $\eta_t$  in a riskless bond with zero interest rate. The trading strategy  $(\theta, \eta)$  gives rise to the following processes.

- Define the discounted value process  $\widetilde{V}_t$ , with respect to the payoff process  $\widehat{\mathbf{S}} = \{(\widehat{S}_t^1, \dots, \widehat{S}_t^N)\}_{t \geq 0}$ , to be

$$\widetilde{V}_t = \sum_{i=1}^N \theta_t^i \widehat{S}_t^i + \eta_t.$$

If the trading strategy  $(\theta, \eta)$  is mean-self-financing, then  $\{\widetilde{V}_t\}_{t \geq 0}$  is a martingale, and vice versa.

- Define the discounted cost process  $\widetilde{C}_t$ , with respect to the payoff process  $\widehat{\mathbf{S}} =$

$\{(\widehat{S}_t^1, \dots, \widehat{S}_t^N)\}_{t \geq 0}$ , to be

$$\begin{aligned} \widetilde{C}_t &= \widetilde{V}_t - \sum_{i=1}^N \int_0^t \theta_s^i d\widehat{S}_s^i, \\ \iff \widetilde{V}_t &= \widetilde{C}_t + \sum_{i=1}^N \int_0^t \theta_s^i d\widehat{S}_s^i. \end{aligned}$$

We note that the processes  $\{\widetilde{V}_t\}_{t \geq 0}$  and  $\{\widetilde{C}_t\}_{t \geq 0}$  defined in this way are martingales, if and only if the trading strategy  $(\theta, \eta)$  is mean-self-financing. Then we can use the hedging scheme in Section 2.6.1, or Elliott and Föllmer (1991).

We use the discounted value process  $\widetilde{V}_t$  to replicate the claim  $H$ 's cum-dividend price process  $\widehat{S}_t = \mathbb{E}^{\mathbb{Q}}(\widetilde{D}_T | \mathcal{F}_t^Z)$ . Since the discounted cost process  $\{\widetilde{C}_t\}_{t \geq 0}$  is a martingale, the trading strategy  $\{(\theta_t^1, \dots, \theta_t^N, \eta_t)\}_{t \geq 0}$  is *mean-self-financing*. The remaining risk is, by definition,  $\mathbb{E}^{\mathbb{Q}}[(\widetilde{C}_T(\theta) - \widetilde{C}_t(\theta))^2 | \mathcal{F}_t^Z]$ .

From the results in Elliott and Föllmer (1991), we know that  $\mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T | \mathcal{F}_t^Z]$  has the Kunita-Watanabe orthogonal representation (decomposition), with respect to the processes  $\{(\widehat{S}_t^1, \dots, \widehat{S}_t^N)\}_{t \geq 0}$ ,

$$\mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T | \mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T] + \Gamma_t + \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i, \quad \text{a.s.} \quad (2.4)$$

for some predictable integrand  $\xi^i$ ,  $i = 1, \dots, N$ . Thus, we can use  $\widetilde{V}_t$  to replicate  $\widehat{S}_t$  of the claim  $H$  as follows.

**Theorem 2.6.1** *Let  $H$  be a derivative (contingent claim) with cumulative discounted dividend stream  $D = \{\widetilde{D}_t\}_{t \geq 0}$  and cum-dividend price  $\widehat{S}_t = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T | \mathcal{F}_t^Z]$ . Let a hedging strategy be defined as*

$$\begin{aligned} \widetilde{V}_t &= \widehat{S}_t = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T] + \Gamma_t + \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i, \\ \theta^i &= \xi^i, \quad i = 1, \dots, N, \\ \eta_t &= \widetilde{V}_t - \sum_{i=1}^N \theta_t^i \widehat{S}_t^i = \widehat{S}_t - \sum_{i=1}^N \theta_t^i \widehat{S}_t^i, \end{aligned} \quad (2.5)$$

where  $\widehat{S}_t$  and  $\xi^i$  are given in the Kunita-Watanabe orthogonal representation (2.4). Then the hedging strategy defined in (2.5) is mean-self-financing, and minimizes the remaining risk  $\mathbb{E}^{\mathbb{Q}}[(\widetilde{C}_T(\theta) - \widetilde{C}_t(\theta))^2 | \mathcal{F}_t^Z]$ .

*Proof.* From the results in Elliott and Föllmer (1991), the cum-dividend price process  $\{\widehat{S}_t\}_{t \geq 0}$  has the Kunita-Watanabe orthogonal representation,

$$\mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T | \mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T] + \Gamma_t + \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i, \quad \text{a.s.} \quad (2.6)$$

for some predictable integrand  $\xi^i$ ,  $i = 1, \dots, N$ , where  $\Gamma = \{\Gamma_t\}_{t \geq 0}$  is a martingale orthogonal to  $\widehat{S}^i$ , for all  $i = 1, \dots, N$ .

The hedging strategy defined in (2.5)

$$\begin{aligned} \widetilde{V}_t &= \widehat{S}_t = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T | \mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T] + \Gamma_t + \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i, \\ \theta^i &= \xi^i, \quad i = 1, \dots, N, \\ \eta_t &= \widetilde{V}_t - \sum_{i=1}^N \theta_t^i \widehat{S}_t^i = \widehat{S}_t - \sum_{i=1}^N \theta_t^i \widehat{S}_t^i, \end{aligned}$$

generates the corresponding cost process

$$\begin{aligned} \widetilde{C}_t &= \widetilde{V}_t - \sum_{i=1}^N \int_0^t \theta_s^i d\widehat{S}_s^i \\ &= \widetilde{V}_t - \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i. \end{aligned}$$

Note that in our hedging strategy defined in (2.5),  $\{\Gamma_t\}_{t \geq 0}$  and each  $\{\widehat{S}_t^i\}_{t \geq 0}$ ,  $1 \leq i \leq N$  is a martingale. Thus, the definition  $\widetilde{V}_t = \mathbb{E}^{\mathbb{Q}}[\widetilde{D}_T] + \Gamma_t + \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i$  implies that  $\{\widetilde{V}_t\}_{t \geq 0}$  is a martingale. Therefore  $\widetilde{C}_t = \widetilde{V}_t - \sum_{i=1}^N \int_0^t \xi_s^i d\widehat{S}_s^i$  implies that the cumulative cost process  $\{\widetilde{C}_t\}_{t \geq 0}$  is a martingale. Hence, by definition, the hedging strategy defined in (2.5) is mean-self-financing.

From the previous notes, since  $\Gamma = \{\Gamma_t\}_{t \geq 0}$  is a martingale orthogonal to  $\widehat{S}^i$ , for all  $i = 1, \dots, N$ , we see that the hedging strategy in (2.5) minimizes the remaining risk  $\mathbb{E}^{\mathbb{Q}}[(\widetilde{C}_T(\theta) - \widetilde{C}_t(\theta))^2 | \mathcal{F}_t^Z]$ .  $\square$

## 2.7 Parameter Estimation

In this section, we shall consider the estimation of the parameters of our model. In our model, the rate matrix  $A := [a_{ij}]_{1 \leq i, j \leq K}$  and the vector of intensity parameters  $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_K)$  are the parameters in the historical measure  $\mathbb{P}$ , and should be calibrated to the times series of historical defaults.

We consider two cases. First, the market observations (market data) are the process of loss states  $Z$ . In this case, we shall use the EM (expectation maximization) algorithm to estimate the parameters. The obtained parameters can be used to calculate the joint default probabilities in the real, physical measure  $\mathbb{P}$ , using Theorems 2.4.1 and 2.4.2.

On the other hand, if we need to also perform pricing of credit risk bearing securities, such as corporate bonds, CDS, and CDO tranche spreads, then in addition to the parameters in  $\mathbb{P}$ , we would also need to know the parameters in the risk-neutral measure  $\mathbb{Q}$ . The estimation of the parameters in the risk-neutral measure  $\mathbb{Q}$  can be performed by a parametric specification of the pricing kernel, which links the historical measure  $\mathbb{P}$  and the pricing measure  $\mathbb{Q}$ . From the time series data of the market price of the securities, we obtain the parameters of the pricing kernel, and therefore obtain the parameters in the pricing measure  $\mathbb{Q}$ , with which we can calculate the securities' prices using the pricing measure  $\mathbb{Q}$ . The parameter estimation for pricing of securities is explored in Section 2.7.2.

### 2.7.1 Parameter Estimation for Default Probabilities from Historical Defaults $Z$

In this section, we consider the parameter estimation for our expected default probability. When the observations of historical loss process  $Z$  are available, we use the EM algorithm to estimate the parameters, in the historical measure  $\mathbb{P}$ .

Recall that  $Z_t^i = \int_0^t \int_E \ell_i \mu^i(du, d\ell) = \ell_i \mathbf{1}_{\{\tau_i \leq t\}}$  is the loss of firm  $i$  by time  $t$ , and  $N_t = \sum_{i=1}^N \int_0^t \int_E \mu^i(du, de)$  is the number of default events of the  $N$  firms up to time  $t$ . We define the following quantities.

1)  $O_t^i$ , the occupation time of the process  $X$  in state  $e_i$  up to time  $t$ , i.e.,

$$O_t^i = \int_0^t \langle X_u, e_i \rangle du.$$

2)  $N_t^{ij}$ , the number of transitions of the process  $X$  from state  $e_i$  to  $e_j$ , up to time  $t$ , i.e.,

$$N_t^{ij} = \int_0^t \langle X_{u-}, e_i \rangle \langle dX_u, e_j \rangle.$$

3)  $G_t^i$ , the level integral for the state  $e_i$  up to time  $t$ , i.e.,

$$G_t^i = \int_0^t \langle X_u, e_i \rangle dN_u.$$

We need to estimate the parameters  $A := [a_{ij}]_{1 \leq i, j \leq K}$  and  $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_K)$ . Let  $\sigma(\cdot)$  denote the conditional expectation  $\mathbb{E}(\cdot | \mathcal{F}_t^Z)$ . We can show that the Maximum Likelihood estimates (MLE)  $\hat{a}_{ij}$  and  $\hat{\lambda}_i$  are

$$\begin{aligned} \hat{a}_{ij} &= \frac{\sigma(N_T^{ij})}{\sigma(O_T^i)} = \frac{\mathbb{E}[N_T^{ij} | \mathcal{F}_T^Z]}{\mathbb{E}[O_T^i | \mathcal{F}_T^Z]} = \frac{\langle \mathbb{E}[N_T^{ij} X_T | \mathcal{F}_T^Z], \mathbf{1} \rangle}{\langle \mathbb{E}[O_T^i X_T | \mathcal{F}_T^Z], \mathbf{1} \rangle} \\ \hat{\lambda}_i &= \frac{\sigma(G_T^i)}{\sigma(O_T^i)} = \frac{\mathbb{E}[G_T^i | \mathcal{F}_T^Z]}{\mathbb{E}[O_T^i | \mathcal{F}_T^Z]} = \frac{\langle \mathbb{E}[G_T^i X_T | \mathcal{F}_T^Z], \mathbf{1} \rangle}{\langle \mathbb{E}[O_T^i X_T | \mathcal{F}_T^Z], \mathbf{1} \rangle} \end{aligned} \quad (2.7)$$

The above estimates involve evaluating  $\sigma(N_T^{ij})$ ,  $\sigma(O_T^i)$  and  $\sigma(G_T^i)$ . In general, it is not possible to compute closed form dynamics for the processes  $\sigma(N^{ij})$ ,  $\sigma(O^i)$  and  $\sigma(G^i)$ . However, it is possible to compute the associated measure-valued quantities  $\sigma(N_t^{ij} X_t)$ ,  $\sigma(O_t^i X_t)$  and  $\sigma(G_t^i X_t)$ , which are vectors in  $\mathbb{R}^K$ . The quantities  $\sigma(N_t^{ij} X_t) = \mathbb{E}[N_t^{ij} X_t | \mathcal{F}_t^Z]$ ,  $\sigma(O_t^i X_t) = \mathbb{E}[O_t^i X_t | \mathcal{F}_t^Z]$ , and  $\sigma(G_t^i X_t) = \mathbb{E}[G_t^i X_t | \mathcal{F}_t^Z]$  can be computed by the following recursive equations.

**Theorem 2.7.1** *The conditional expectations  $\mathbb{E}[N_t^{\ell m} X_t | \mathcal{F}_t^Z]$ ,  $\mathbb{E}[O_t^\ell X_t | \mathcal{F}_t^Z]$ , and  $\mathbb{E}[G_t^\ell X_t | \mathcal{F}_t^Z]$  satisfy*

$$\sigma(G_t^\ell X_t) = \int_0^t A \sigma(G_u^\ell X_u) du + \sum_{i=1}^N \int_0^t \text{diag}(H(u) - 1) \sigma(G_{u-}^\ell X_{u-}) dN_u^i$$

$$\begin{aligned}
& - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda - 1) \sigma(G_u^\ell X_u) du \\
& + \sum_{i=1}^N \int_0^t \langle H(u), e_i \rangle \langle q_{u-}, e_\ell \rangle dN_u^i e_\ell - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \langle \lambda - 1, e_i \rangle \langle q_u, e_\ell \rangle du e_\ell, \quad (2.8)
\end{aligned}$$

$$\begin{aligned}
\sigma(N_t^{\ell m} X_t) &= \int_0^t A \sigma(N_u^{\ell m} X_u) du + \sum_{i=1}^N \int_0^t \text{diag}(H(u) - 1) \sigma(N_{u-}^{\ell m} X_{u-}) dN_u^i \\
& - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda - 1) \sigma(N_u^{\ell m} X_u) du + \int_0^t \langle q_{u-}, e_\ell \rangle \langle A e_\ell, e_m \rangle du e_\ell, \quad (2.9)
\end{aligned}$$

$$\begin{aligned}
\sigma(O_t^\ell X_t) &= \int_0^t A \sigma(O_u^\ell X_u) du + \sum_{i=1}^N \int_0^t \text{diag}(H(u) - 1) \sigma(O_{u-}^\ell X_{u-}) dN_u^i \\
& - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda - 1) \sigma(O_u^\ell X_u) du + \int_0^t \langle q_u, e_\ell \rangle du e_\ell. \quad (2.10)
\end{aligned}$$

*Proof.* See Appendix 4. □

We can compute the estimates  $\hat{a}_{ij}$  and  $\hat{\lambda}_i$  by the following EM algorithm.

- I:** Set the initial values for  $\hat{a}_{ij}(0)$  and  $\hat{\lambda}_i(0)$ .
- II:** Compute the updated ML estimates  $\hat{a}_{ij}(k+1)$  and  $\hat{\lambda}_i(k+1)$ , using (2.7), (2.8)–(2.10) and the market observations for  $Z$ .
- III:** Stop if  $|\hat{a}_{ij}(k+1) - \hat{a}_{ij}(k)|$  and  $|\hat{\lambda}_i(k+1) - \hat{\lambda}_i(k)|$  arrive at the desirable precision; otherwise, continue Step II.

## 2.7.2 Joint Parameter Estimation for the Pricing of Credit Securities

In this section, we propose a framework which is able to jointly determine default probabilities and security prices. The joint estimation for the parameters uses time-series data of both historical defaults and security price.

From Section 2.7.1, we have the parameters  $A^{\mathbb{P}}$  and  $\lambda^{\mathbb{P}}$  for the historical distribution. In order to obtain the parameters for the pricing, we need parameters  $A^{\mathbb{Q}}$  and  $\lambda^{\mathbb{Q}}$  in the risk-neutral measure  $\mathbb{Q}$ . For simplicity we could suppose  $A^{\mathbb{Q}} = A^{\mathbb{P}}$ . We obtain  $\lambda^{\mathbb{Q}}$  through a

parametric specification of the pricing kernel  $M_t$ . In general, we know that  $M_t = e^{-\int_0^t r_u du} \Lambda_t$ , where  $\Lambda_t = d\mathbb{Q}/d\mathbb{P}$  is the Radon-Nikodym derivative, i.e.,

$$\Lambda_t = \exp \left( - \sum_{i=1}^N \int_0^t \int_E (h_u^i(y) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f_i^{\mathbb{P}}(y) dy du + \sum_{i=1}^N \int_0^t \int_E \log h_u^i(y) \mu^i(du, dy) \right).$$

Here

$$h_u^i(y) := \frac{\lambda_{iu}^{\mathbb{Q}} f_i^{\mathbb{Q}}(y)}{\lambda_{iu}^{\mathbb{P}} f_i^{\mathbb{P}}(y)} = \sum_{k=1}^K \langle X_{u-}, e_k \rangle \frac{\lambda_{ik}^{\mathbb{Q}} f_i^{\mathbb{Q}}(y)}{\lambda_{ik}^{\mathbb{P}} f_i^{\mathbb{P}}(y)} = \sum_{k=1}^K \langle X_{u-}, e_k \rangle h_k^i(y).$$

Using Itô's rule, we have

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \int_E \Lambda_{u-} (h_u^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} \lambda_{iu}^{\mathbb{P}} f_i^{\mathbb{P}}(y) dy du),$$

where  $\mu^i(dy, ds) = \delta_{y\tau}(dy) \delta_\tau(ds)$  is the random measure associated with the jump process  $Z^i$ .

Suppose we have time-series data on both historical defaults and the market price of the credit derivatives (e.g., credit spread of bonds). We can obtain  $A^{\mathbb{P}}$  and  $\lambda^{\mathbb{P}}$  from the historical defaults, as indicated in Section 2.7.1. Then we could parameterize a rich and still tractable relationship between the risk-neutral parameters  $\lambda^{\mathbb{Q}}$  and  $\lambda^{\mathbb{P}}$  as follows. We set  $\lambda_t^{\mathbb{Q}} = a\lambda_t^{\mathbb{P}} + bX_t$ . For simplicity, we can set  $f_i^{\mathbb{Q}}(y) = f_i^{\mathbb{P}}(y)$ , namely, the distribution of random jump sizes of the process  $Z^i$  in  $\mathbb{Q}$  is the same as in  $\mathbb{P}$ . This means that we are concerned more with the default event risk, and less with the recovery risk. A similar strategy is adapted in Pan (2002). More general affine specifications of  $\lambda^{\mathbb{Q}}$  from  $\lambda^{\mathbb{P}}$  are discussed in Duffie and Singleton (2003).

From the time-series data of the market security prices (e.g., credit spread of bonds), we can use simulation based estimation, e.g., the Simulated Moment Method, to obtain estimates of  $a$  and  $b$ , and thus  $\lambda^{\mathbb{Q}}$ . With  $\lambda^{\mathbb{Q}}$  we have the pricing kernel (stochastic discount factor) and therefore the risk-neutral measure  $\mathbb{Q}$ , where the pricing can be performed.

## 2.8 Concluding Remarks

We have investigated models for pricing credit derivatives. Our models naturally allow for both frailty and default contagion.

We consider models with various number of underlying Markov processes. We obtain the estimates of  $X_t$  and the pricing formulas for general credit securities when the model has one, two, and three underlying processes, respectively. We obtain **closed-form** solutions for the joint default probabilities. We also give dynamic hedging for the credit securities discussed in our model. Our model allows for joint estimate of parameters from both historical defaults and securities prices.

Our method *can be generalized* to estimate  $X, X^1, \dots, X^N$  in our general, ideal model with  $N + 1$  chains in Section 3.2.1, from the observations process  $Z$ . However, in such a general model, the resulting calculations and simulations become complicated.

In this paper, for simplicity, we only investigate parameter estimation for the simple model with just one underlying Markov chain. This is done by the EM (expectation maximization) algorithm. Our approach can be generalized to obtain the estimates of the parameters in more complicated models, for example, the general model with  $N + 1$  underlying Markov chains  $X, X^1, X^2, \dots, X^N$ .

## Appendix 1

Proof of Theorem 2.3.1

For each  $k = 1, 2, \dots, K$ ,  $f_k^i(y)$  is the probability density function for the random jump size  $y^i = Z_u^i - Z_{u-}^i = \mathbf{1}_{\{u=\tau_i\}} \ell_u^i$ , when  $X_{u-} = e_k$ . Under the probability measure  $\mathbb{P}$ , the compensator of the random measure  $\mu^i(du, dy)$  is

$$\nu^i(du, dy|X_{u-}) := \sum_{k=1}^K \langle X_{u-}, e_k \rangle \mathbf{1}_{\{u \leq \tau_i\}} \lambda_k^i f_k^i(y) du dy.$$



Consequently, under  $\mathbb{P}$ , with necessary integrability conditions,

$$Z_t^i - \int_0^t \int_E y \nu^i(du, dy | X_{u-}) = \int_0^t \int_E y(\mu^i(du, dy) - \nu^i(du, dy | X_{u-}))$$

is a martingale.

Suppose we have a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq K$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y)$ , independent of the hidden state  $X$ , for the loss sizes. Then, the compensator  $\bar{\nu}^i$  of  $\mu^i$  under  $\bar{\mathbb{P}}$  is

$$\bar{\nu}^i(du, dy) := \mathbf{1}_{\{u \leq \tau_i\}} f(y) du dy.$$

Consequently, under  $\bar{\mathbb{P}}$ , with necessary integrability conditions,

$$Z_t^i - \int_0^t \int_E y \bar{\nu}^i(du, dy | X_{u-}) = \int_0^t \int_E y \mu^i(du, dy) - \int_0^t \int_E y \mathbf{1}_{\{u \leq \tau_i\}} f(y) du dy$$

is a martingale. We start with  $\bar{\mathbb{P}}$ . Using Girsanov's Theorem for jump processes, define a process  $\Lambda_t$  by

$$\Lambda_t = \prod_{i=1}^N \exp \left( - \int_0^t \int_E \sum_{k=1}^K \langle X_{u-}, e_k \rangle (h_k^i(y) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y) du dy + \int_0^t \int_E \sum_{k=1}^K \langle X_{u-}, e_k \rangle \log h_k^i(y) \mu^i(du, dy) \right) \quad (2.11)$$

$$= \exp \left( - \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle X_{u-}, e_k \rangle (h_k^i(y) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y) du dy + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle X_{u-}, e_k \rangle \log h_k^i(y) \mu^i(du, dy) \right)$$

Using Itô's rule, we also have

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \int_E \Lambda_{u-} \sum_{k=1}^K \langle X_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du),$$

or

$$d(\Lambda_t) = \Lambda_{t-} \sum_{i=1}^N \int_E \sum_{k=1}^K \langle X_{t-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{t \leq \tau_i\}} f(y) dy dt).$$

Since

$$d(\Lambda_t X_t) = d(\Lambda_t) X_t + \Lambda_t d(X_t)$$

$$= \Lambda_{t-} \left( \sum_{i=1}^N \int_E \sum_{k=1}^K \langle X_{t-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy dt) \right) X_{t-} + \Lambda_{t-} (AX_t + dM_t),$$

we have

$$\begin{aligned} \Lambda_t X_t &= \Lambda_0 X_0 + \sum_{i=1}^N \int_0^t \Lambda_{u-} X_{u-} \int_E \sum_{k=1}^K \langle X_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) \\ &\quad + \int_0^t \Lambda_u A X_u du + \int_0^t \Lambda_{u-} dM_u. \end{aligned}$$

Write  $q_t := \overline{\mathbb{E}}[\Lambda_t X_t | \mathcal{F}_t^Z]$ . Then, a version of Bayes' Theorem gives

$$\mathbb{E}[X_t | \mathcal{F}_t^Z] = \frac{\overline{\mathbb{E}}[\Lambda_t X_t | \mathcal{F}_t^Z]}{\overline{\mathbb{E}}[\Lambda_t | \mathcal{F}_t^Z]} = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}.$$

Conditioning on  $\mathcal{F}_t^Z$  under  $\overline{\mathbb{P}}$ , using the Fubini Theorem in Hajek and Wong (1985), we have

$$q_t = q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle q_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) e_k$$

Write  $H_k^i(u) := h_k^i(y_u)$ . Then

$$\begin{aligned} q_t &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle q_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) e_k \\ &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle q_{u-}, e_k \rangle (h_k^i(y) - 1) \mu^i(dy, du) e_k \\ &\quad - \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle q_{u-}, e_k \rangle (h_k^i(y) - 1) (\mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) e_k \\ &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \text{diag}(H_1^i(u) - 1, \dots, H_K^i(u) - 1) q_{u-} \times \int_E \mu^i(dy, du) \\ &\quad - \sum_{i=1}^N \int_0^t \text{diag}(\lambda_1^i - 1, \dots, \lambda_K^i - 1) \mathbf{1}_{\{u \leq \tau_i\}} q_u du \\ &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \text{diag}(H_1^i(u) - 1, \dots, H_K^i(u) - 1) q_{u-} dN_u^i \\ &\quad - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda_1^i - 1, \dots, \lambda_K^i - 1) q_u du. \end{aligned}$$

□

## Appendix 2

Proof of Theorem 2.3.2

Recall that for each  $k = 1, 2, \dots, K$ , and  $m = 1, 2, \dots, M$ ,  $f_{km}^i(y)$  is the probability density function for the random jump size  $y^i = \mathbf{1}_{\{u=\tau_i\}}\ell_u^i$ , when  $X_{u-} = e_k$  and  $Y_{u-} = \epsilon_m$ . Then, under the probability distribution  $\mathbb{P}$ , the compensator of the random measure  $\mu^i(du, dy)$  is

$$\nu^i(du, dy^i | X_{u-}, Y_{u-}) := \sum_{m=1}^M \sum_{k=1}^K \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (\alpha_k^i + \beta_m^i) \mathbf{1}_{\{u \leq \tau_i\}} f_{km}^i(y^i) du dy^i.$$

Consequently, under  $\mathbb{P}$ , with necessary integrability conditions,

$$Z_t^i - \int_0^t \int_E y^i \nu^i(du, dy^i | X_{u-}, Y_{u-}) = \int_0^t \int_E y^i (\mu^i(du, dy^i) - \nu^i(du, dy^i | X_{u-}, Y_{u-}))$$

is a martingale.

Suppose we have a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq N$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y^i)$ , independent of the hidden states  $X_t$  and  $Y_t$ , for the loss sizes. Then, the compensator  $\bar{\nu}^i$  of  $\mu^i$  under  $\bar{\mathbb{P}}$  is

$$\bar{\nu}^i(du, dy) := \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) du dy^i.$$

Consequently, under  $\bar{\mathbb{P}}$ , with necessary integrability conditions,

$$Z_t^i - \int_0^t \int_E y^i \bar{\nu}^i(du, dy^i | X_{u-}, Y_{u-}) = \int_0^t \int_E y^i \mu^i(du, dy^i) - \int_0^t \int_E y^i \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) du dy^i$$

is a martingale.

We start with  $\bar{\mathbb{P}}$ . Using Girsanov's Theorem for jump processes, define a process  $\Lambda_t$  by

$$\begin{aligned} \Lambda_t = \prod_{i=1}^N \exp \left( - \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (h_{km}^i(y^i) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) du dy^i \right. \\ \left. + \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle \log h_{km}^i(y^i) \mu^i(du, dy^i) \right) \end{aligned} \quad (2.12)$$

$$\begin{aligned}
&= \exp \left( - \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (h_{km}^i(y^i) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) du dy^i \right. \\
&\quad \left. + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle \log h_{km}^i(y^i) \mu^i(du, dy^i) \right)
\end{aligned}$$

Using Itô's rule, we also have

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \int_E \Lambda_{u-} \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (h_{km}^i(y^i) - 1) (\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du),$$

or

$$d(\Lambda_t) = \Lambda_{t-} \sum_{i=1}^N \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (h_{km}^i(y^i) - 1) (\mu^i(dy^i, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i dt),$$

Let  $\otimes$  denote the tensor product (outer product) of two vectors, i.e.  $X \otimes Y := X(Y)^\top$ . Since

$$\begin{aligned}
d(\Lambda_t X_t \otimes Y_t) &= (d\Lambda_t) X_t \otimes Y_t + \Lambda_t (dX_t) \otimes Y_t + \Lambda_t X_t \otimes (dY_t) \\
&= \Lambda_{t-} \left( \sum_{i=1}^N \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (h_{km}^i(y^i) - 1) (\mu^i(dy^i, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i dt) \right) \\
&\quad \times (X_t \otimes Y_t) + \Lambda_t (A^1 X_t dt + dM_t^1) \otimes Y_t + \Lambda_t X_t \otimes (A^1 Y_t dt + dM_t^2).
\end{aligned}$$

we have

$$\begin{aligned}
\Lambda_t X_t \otimes Y_t &= \Lambda_0 X_0 \otimes Y_0 + \\
&\quad \sum_{i=1}^N \int_0^t \Lambda_{u-} (X_{u-} \otimes Y_{u-}) \int_E \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}, \epsilon_m \rangle (h_{km}^i(y^i) - 1) (\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) \\
&\quad + \int_0^t \Lambda_u (A^1 X_u du + dM_u^1) \otimes Y_u + \Lambda_u X_u \otimes (A^2 Y_u du + dM_u^2).
\end{aligned}$$

Write  $q_t := \overline{\mathbb{E}}[\Lambda_t X_t \otimes Y_t | \mathcal{F}_t^M]$ . Then, a version of Bayes' Theorem gives

$$\mathbb{E}[X_t \otimes Y_t | \mathcal{F}_t^M] = \frac{\overline{\mathbb{E}}[\Lambda_t X_t \otimes Y_t | \mathcal{F}_t^M]}{\overline{\mathbb{E}}[\Lambda_t | \mathcal{F}_t^M]} = \frac{q_t}{(\mathbf{1}_K)^\top q_t \mathbf{1}_M}.$$

Conditioning on  $\mathcal{F}_t^Z$  under  $\overline{\mathbb{P}}$ , we have

$$q_t = q_0 + \int_0^t A^1 q_u du + \int_0^t q_u (A^2)^\top du$$

$$+ \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M (e_k^\top q_{u-\epsilon_m})(h_{km}^i(y^i) - 1)(\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) e_k \otimes \epsilon_m.$$

Write  $H_{km}^i(u) := h_{km}^i(y_u)$ . Then

$$\begin{aligned} q_t &= q_0 + \int_0^t A^1 q_u du + \int_0^t q_u (A^2)^\top du + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M (e_k^\top q_{u-\epsilon_m})(h_{km}^i(y^i) - 1) \mu^i(dy^i, du) e_k \otimes \epsilon_m \\ &\quad - \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M (e_k^\top q_{u-\epsilon_m})(h_{km}^i(y^i) - 1) (\mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) e_k \otimes \epsilon_m \\ &= q_0 + \int_0^t A^1 q_u du + \int_0^t q_u (A^2)^\top du + \sum_{i=1}^N \int_0^t \begin{pmatrix} H_{11}^i(u) - 1 & \dots & H_{1M}^i(u) - 1 \\ \vdots & \ddots & \vdots \\ H_{K1}^i(u) - 1 & \dots & H_{KM}^i(u) - 1 \end{pmatrix} \circ q_{u-} \cdot \left( \int_E \mu^i(dy^i, du) \right) \\ &\quad - \sum_{i=1}^N \int_0^t \begin{pmatrix} (\alpha_1^i + \beta_1^i) - 1 & \dots & (\alpha_1^i + \beta_M^i) - 1 \\ \vdots & \ddots & \vdots \\ (\alpha_K^i + \beta_1^i) - 1 & \dots & (\alpha_K^i + \beta_M^i) - 1 \end{pmatrix} \circ (\mathbf{1}_{\{u \leq \tau_i\}} q_u) du \\ &= q_0 + \int_0^t A^1 q_u du + \int_0^t q_u (A^2)^\top du + \sum_{i=1}^N \int_0^t \begin{pmatrix} H_{11}^i(u) - 1 & \dots & H_{1M}^i(u) - 1 \\ \vdots & \ddots & \vdots \\ H_{K1}^i(u) - 1 & \dots & H_{KM}^i(u) - 1 \end{pmatrix} \circ q_{u-} dN_u^i \\ &\quad - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \begin{pmatrix} (\alpha_1^i + \beta_1^i) - 1 & \dots & (\alpha_1^i + \beta_M^i) - 1 \\ \vdots & \ddots & \vdots \\ (\alpha_K^i + \beta_1^i) - 1 & \dots & (\alpha_K^i + \beta_M^i) - 1 \end{pmatrix} \circ q_u du, \end{aligned}$$

where the operation  $\circ$  denotes the Hadamard product (entrywise product) of two matrices.

□

## Appendix 3

### Proof of Theorem 2.3.3

For each  $k = 1, 2, \dots, K$ ,  $m = 1, 2, \dots, M$  and  $j = 1, 2, \dots, J$ ,  $f_{kmj}^i(y)$  is the probability density function for the random jump size  $y^i = \mathbf{1}_{\{u=\tau_i\}} \ell_u^i$ , when  $X_{u-} = e_k$ ,  $Y_{u-}^1 = \epsilon_m$  and

$Y_{u-}^2 = \delta_j$ . We know that, under the probability distribution  $\mathbb{P}$ , the compensator of the random measure  $\mu^i(du, dy)$  is

$$\nu^i(du, dy^i | X_{u-}, Y_{u-}^1, Y_{u-}^2) := \sum_{j=1}^J \sum_{k=1}^K \sum_{m=1}^M \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (\alpha_k^i + \beta_m^i + \zeta_j^i) \mathbf{1}_{\{u \leq \tau_i\}} f_{kmj}^i(y^i) dudy^i.$$

Consequently, under  $\mathbb{P}$ , with necessary integrability conditions,

$$Z_t^i - \int_0^t \int_E y^i \nu^i(du, dy^i | X_{u-}, Y_{u-}^1, Y_{u-}^2) = \int_0^t \int_E y^i (\mu^i(du, dy^i) - \nu^i(du, dy^i | X_{u-}, Y_{u-}^1, Y_{u-}^2))$$

is a martingale.

Suppose we have a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq N$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y^i)$ , independent of the hidden states  $X_t$  and  $Y_t^1$ , for the loss sizes. Then, the compensator  $\bar{\nu}^i$  of  $\mu^i$  under  $\bar{\mathbb{P}}$  is

$$\bar{\nu}^i(du, dy) := \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dudy^i.$$

Consequently, under  $\bar{\mathbb{P}}$ , with necessary integrability conditions,

$$Z_t^i - \int_0^t \int_E y^i \bar{\nu}^i(du, dy^i | X_{u-}, Y_{u-}^1, Y_{u-}^2) = \int_0^t \int_E y^i \mu^i(du, dy^i) - \int_0^t \int_E y^i \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dudy^i$$

is a martingale.

We start with  $\bar{\mathbb{P}}$ . Using Girsanov's Theorem for jump processes, define a process  $\Lambda_t$  by

$$\begin{aligned} \Lambda_t &= \prod_{i=1}^N \exp \left( - \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (h_{kmj}^i(y^i) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dudy^i \right. \\ &\quad \left. + \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle \log h_{kmj}^i(y^i) \mu^i(du, dy^i) \right) \\ &= \exp \left( - \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (h_{kmj}^i(y^i) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dudy^i \right. \\ &\quad \left. + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle \log h_{kmj}^i(y^i) \mu^i(du, dy^i) \right). \end{aligned} \tag{2.13}$$

Using Itô's rule, we also have

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \int_E \Lambda_{u-} \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du),$$

or

$$d(\Lambda_t) = \Lambda_{t-} \sum_{i=1}^N \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mu^i(dy^i, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i dt).$$

Let  $\odot$  denote the Kronecker product of two matrices. The Kronecker product has the following property

$$(A \odot B)(P \odot Q) = (AP) \odot (BQ).$$

Since

$$\begin{aligned} & d(\Lambda_t X_t \odot Y_t^1 \odot Y_t^2) \\ &= (d\Lambda_t) X_t \odot Y_t^1 \odot Y_t^2 + \Lambda_t (dX_t) \odot Y_t^1 \odot Y_t^2 + \Lambda_t X_t \odot (dY_t^1) \odot Y_t^2 + \Lambda_t X_t \odot Y_t^1 \odot (dY_t^2) \\ &= (X_t \odot Y_t^1 \odot Y_t^2) \Lambda_t \\ &\quad \times \left( \sum_{i=1}^N \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mu^i(dy^i, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i dt) \right) \\ &\quad + \Lambda_t (A^1 X_t dt + dM_t^1) \odot Y_t^1 \odot Y_t^2 + \Lambda_t X_t \odot (A^2 Y_t^1 dt + dM_t^2) \odot Y_t^2 + \Lambda_t X_t \odot Y_t^1 \odot (A^3 Y_t^2 dt + dM_t^3). \end{aligned}$$

we have

$$\begin{aligned} \Lambda_t X_t \odot Y_t^1 \odot Y_t^2 &= \Lambda_0 X_0 \odot Y_0^1 \odot Y_0^2 + \sum_{i=1}^N \int_0^t \Lambda_{u-} (X_{u-} \odot Y_{u-}^1 \odot Y_{u-}^2) \\ &\quad \times \int_E \sum_{k,m,j} \langle X_{u-}, e_k \rangle \langle Y_{u-}^1, \epsilon_m \rangle \langle Y_{u-}^2, \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) \\ &\quad + \int_0^t \Lambda_u (A^1 X_u du + dM_u^1) \odot Y_u^1 \odot Y_u^2 + \Lambda_u X_u \odot (A^2 Y_u^1 du + dM_u^2) \odot Y_u^2 \\ &\quad + \Lambda_u X_u \odot Y_u^1 \odot (A^3 Y_u^2 du + dM_u^3). \end{aligned}$$

Write  $q_t := \overline{\mathbb{E}}[\Lambda_t X_t \odot Y_t^1 \odot Y_t^2 | \mathcal{F}_t^M]$ . Then, a version of Bayes' Theorem gives

$$\mathbb{E}[X_t \odot Y_t^1 \odot Y_t^2 | \mathcal{F}_t^M] = \frac{\overline{\mathbb{E}}[\Lambda_t X_t \odot Y_t^1 \odot Y_t^2 | \mathcal{F}_t^M]}{\overline{\mathbb{E}}[\Lambda_t | \mathcal{F}_t^M]} = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}.$$

Conditioning on  $\mathcal{F}_t^M$  under  $\bar{\mathbb{P}}$ , we have

$$\begin{aligned} q_t &= q_0 + \int_0^t (A^1 \odot I_M \odot I_J) q_u du + \int_0^t (I_K \odot A^2 \odot I_J) q_u du + \int_0^t (I_K \odot I_M \odot A^3) q_u du \\ &\quad + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle q_{u-}, e_k \odot \epsilon_m \odot \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) e_k \odot \epsilon_m \odot \delta_j. \end{aligned}$$

Write  $H_{kmj}^i(u) := h_{kmj}^i(y_u)$ . Then

$$\begin{aligned} q_t &= q_0 + \int_0^t (A^1 \odot I_M \odot I_J) q_u du + \int_0^t (I_K \odot A^2 \odot I_J) q_u du + \int_0^t (I_K \odot I_M \odot A^3) q_u du \\ &\quad + \sum_{i=1}^N \int_0^t \int_E \sum_{k,m,j} \langle q_{u-}, e_k \odot \epsilon_m \odot \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mu^i(dy^i, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) e_k \odot \epsilon_m \odot \delta_j \\ &= q_0 + \int_0^t (A^1 \odot I_M \odot I_J) q_u du + \int_0^t (I_K \odot A^2 \odot I_J) q_u du + \int_0^t (I_K \odot I_M \odot A^3) q_u du \\ &\quad + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle q_{u-}, e_k \odot \epsilon_m \odot \delta_j \rangle (h_{kmj}^i(y^i) - 1) \mu^i(dy^i, du) e_k \odot \epsilon_m \odot \delta_j \\ &\quad - \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \langle q_{u-}, e_k \odot \epsilon_m \odot \delta_j \rangle (h_{kmj}^i(y^i) - 1) (\mathbf{1}_{\{u \leq \tau_i\}} f(y^i) dy^i du) e_k \odot \epsilon_m \odot \delta_j \\ &= q_0 + \int_0^t (A^1 \odot I_M \odot I_J) q_u du + \int_0^t (I_K \odot A^2 \odot I_J) q_u du + \int_0^t (I_K \odot I_M \odot A^3) q_u du \\ &\quad + \sum_{i=1}^N \int_0^t (H_{111}^i(u) - 1, \dots, H_{11J}^i(u) - 1, \dots, H_{KMJ}^i(u) - 1)^\top \circ q_{u-} \cdot \left( \int_E \mu^i(dy^i, du) \right) \\ &\quad - \sum_{i=1}^N \int_0^t ((\alpha_1^i + \beta_1^i + \zeta_1^i) - 1, \dots, (\alpha_1^i + \beta_1^i + \zeta_J^i) - 1, \dots, (\alpha_K^i + \beta_M^i + \zeta_J^i) - 1)^\top \circ (\mathbf{1}_{\{u \leq \tau_i\}} q_u) du \\ &= q_0 + \int_0^t (A^1 \odot I_M \odot I_J) q_u du + \int_0^t (I_K \odot A^2 \odot I_J) q_u du + \int_0^t (I_K \odot I_M \odot A^3) q_u du \\ &\quad + \sum_{i=1}^N \int_0^t (H_{111}^i(u) - 1, \dots, H_{11J}^i(u) - 1, \dots, H_{KMJ}^i(u) - 1)^\top \circ q_{u-} \cdot dN_u^i \\ &\quad - \sum_{i=1}^N \int_0^{t \wedge \tau_i} ((\alpha_1^i + \beta_1^i + \zeta_1^i) - 1, \dots, (\alpha_1^i + \beta_1^i + \zeta_J^i) - 1, \dots, (\alpha_K^i + \beta_M^i + \zeta_J^i) - 1)^\top \circ q_u du, \end{aligned}$$

where the operation  $\circ$  denotes the Hadamard product (entrywise product) of two matrices.

□



## Appendix 4

Proof of Theorem 2.7.1

Recall that

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \int_E \Lambda_{u-} \sum_{k=1}^K \langle X_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du),$$

or

$$d(\Lambda_t) = \Lambda_{t-} \sum_{i=1}^N \int_E \sum_{k=1}^K \langle X_t, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy dt).$$

Using the product rule, the process  $\Lambda G^\ell X$  is

$$\begin{aligned} \Lambda_t G_t^\ell X_t &= \int_0^t \Lambda_{u-} X_u \langle X_u, e_\ell \rangle dN_u + \int_0^t \Lambda_u G_u^\ell A X_u du + \int_0^t \Lambda_u G_u^\ell dM_u \\ &\quad + \int_0^t G_u^\ell X_u \Lambda_{u-} \left( \sum_{i=1}^N \int_E \sum_{k=1}^K \langle X_u, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy dt) \right) \\ &\quad + \sum_{i=1}^N \int_0^t \int_E X_u \Lambda_{u-} \langle X_u, e_\ell \rangle (h_\ell^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy dt) \end{aligned}$$

Taking the conditional expectation under  $\bar{\mathbb{P}}$ ,

$$\begin{aligned} \sigma(G_t^\ell X_t) &= \int_0^t \langle q_{u-}, e_\ell \rangle dN_u e_\ell + \int_0^t A \sigma(G_u^\ell X_u) du + \sum_{i=1}^N \int_0^t \text{diag}(H(u) - 1) \sigma(G_{u-}^\ell X_{u-}) dN_u^i \\ &\quad - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda - 1) \sigma(G_u^\ell X_u) du \\ &\quad + \sum_{i=1}^N \int_0^t \langle H(u) - 1, e_i \rangle \langle q_{u-}, e_\ell \rangle dN_u^i e_\ell - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \langle \lambda - 1, e_i \rangle \langle q_u, e_\ell \rangle du e_\ell \\ &= \int_0^t A \sigma(G_u^\ell X_u) du + \sum_{i=1}^N \int_0^t \text{diag}(H(u) - 1) \sigma(G_{u-}^\ell X_{u-}) dN_u^i \\ &\quad - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda - 1) \sigma(G_u^\ell X_u) du \\ &\quad + \sum_{i=1}^N \int_0^t \langle H(u), e_i \rangle \langle q_{u-}, e_\ell \rangle dN_u^i e_\ell - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \langle \lambda - 1, e_i \rangle \langle q_u, e_\ell \rangle du e_\ell, \end{aligned}$$

Similarly, we can get the recursive equations for  $\sigma(N_t^{ij} X_t)$  and  $\sigma(O_t^i X_t)$ .  $\square$

## Chapter 3

# Credit Risk and Contagion via Self-Exciting Default Intensity

### 3.1 Introduction

Modeling credit risk is one of the central topics in financial research. There are generally two major approaches to model credit risk: the structural model and the reduced-form model. The structural (firm value based) approach, initiated by Merton (1974) and Black and Cox (1976), considers the relationship between a firm's asset value and the firm's default event. The default of the firm is triggered when the asset value falls below a default level. On the other hand, the reduced-form approach, introduced by Jarrow and Turnbull (1995), Jarrow, Lando and Turnbull (1997), Duffie and Singleton (1997, 1999), etc., treats defaults as exogenous events and models the arrivals of defaults using doubly stochastic Poisson point processes.

Credit risk models with incomplete information have been considered by various researchers. Recent literature for both structural and reduced-form models includes Collin-Dufresne, Goldstein and Helwege (2003), Schönbucher (2003), Duffie and Lando (2001), Duffie, Eckner, Horel and Saita (2009), Frey and Runggaldier (2010), and Frey and Schmidt (2012), etc. The reduced-form models in these papers accommodate some common features: for example, default intensities are driven by an underlying (state) process  $X$ ; conditional on the current information, the default times are assumed to be random times; investors have access to partial information consisting of observations of the market, as well as observable economic covariates. Collin-Dufresne *et al.* (2003) and Schönbucher (2003) model the underlying factor  $X$  as a static random vector. Moreover, it is pointed out in Collin-Dufresne *et al.*

(2003) and Schönbucher (2003) that the event that some firm has defaulted causes an update of the distribution of the factor process, and hence causes a jump in the default intensity of other surviving firms. Therefore, the successive updating of the distribution of the state process in reaction to incoming default observations is likely to give rise to contagion effects. An important example for default contagion is evidenced by the default of Lehman Brothers in 2008 and the subsequent market crash. Duffie *et al.* (2009) model the state process  $X$  as an Ornstein–Uhlenbeck process. Furthermore, the empirical analysis in Duffie *et al.* (2009) suggests that, in addition to the observable economic covariates, an unobservable process driving default intensities is needed to account for the clustering of defaults in historical data.

In various previous reduced-form models, the default intensity is usually modeled as a *doubly stochastic* process. In a doubly stochastic framework, default arrivals are assumed to be conditionally Poisson given the paths of the observable and/or latent risk factors. However, Das et al. (2007) conduct empirical analyses to test the doubly-stochastic assumption that firms' defaults only depend on these observable and/or latent factors. Das et al. (2007) present evidence that the U.S. corporations' historical defaults are inconsistent with the doubly stochastic model, and find that there is too much default clustering relative to that predicted by the doubly stochastic model. Furthermore, Azizpour, Giesecke and Schwenkler (2014) conduct extensive empirical research, based on a data set with significantly longer sample period (from 1970 to 2012) containing the recent financial crisis and larger cross-section set including both industrial and financial firms. Azizpour, Giesecke and Schwenkler (2014) firmly reject the hypothesis that the economy-wide default rate is influenced only by the risk factors. Their empirical results indicate that the doubly-stochastic models that have been previously widely used may understate the risk of large losses from defaults, especially during clustering periods such as 2001-2002 and 2007-2009. Moreover, the authors find strong evidence that the conditional default rate depends on past failures, regardless of

whether frailty factor is present.

Another important empirical finding in Azizpour, Giesecke and Schenkler (2014) is that firms usually have joint exposure to both contagion and frailty effects, in addition to the observable macro-economic risk factors. For example, empirical evidences in Azizpour, Giesecke and Schenkler (2014) indicate that contagion effect plays a more prominent role for explaining the default clusters than the frailty factor; the in-sample fit of a model including the contagion effect outperforms a model without that effect; a model ignoring the contagion effect tends to overstate the frailty effect. Out-of-sample tests in Azizpour, Giesecke and Schenkler (2014) indicate that models with contagion and frailty clustering sources perform much better on forecasting defaults than models where firms only have exposure to the observable macro-economic risk factors; a model without the contagion effect generates excessively high and volatile forecasts, while a model without the frailty effect tends to understate the forecasts. Only a model with all three sources of clustering provides accurate and sensible out-of-sample forecasts of correlated defaults.

Our model is mainly motivated by the empirical findings presented in Azizpour, Giesecke and Schenkler (2014) that default rate depends not only on the risk factors, but also on the past defaults. Roughly speaking, differently to the previous contributions, we consider a reduced-form model where the intensity  $\lambda$  is “self-exciting” in the sense that  $\lambda$  varies not only with the underlying state process  $X$ , but also depends on the previous default history of all the firms. Our model with the “self-exciting” default intensity and the underlying Markov state process allows for both contagion and frailty effects. More precisely, the default intensity in our model depends on the path of the past default history. Thus the contagion effect arises from the past defaults are modeled, in a sense, through a so-called “mean-field” approach. Recent research using a similar “mean-field” approach to model correlated default risk includes Cvitanić, Ma and Zhang (2010), and Giesecke, Spiliopoulos and R. Sowers (2013). In Cvitanić, Ma and Zhang (2010), the default intensity of a firm depends

on the percentage loss of the portfolio. In Giesecke, Spiliopoulos and R. Sowers (2013), the default intensity of a firm depends on the default rate of the portfolio.

In our model, the states of the Markov Chains represent different economic environments. The switching of the states of the economy can be attributed to structural changes in macro-economic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities.

In practice, investors can observe the number of defaults and individual default loss amounts over a given time period. However, the underlying states of the economy, or the regime-switching process, are not directly observable. Therefore, we further assume that the background state processes are *Hidden Markov chains*, namely the states of the chains are not directly observable. Rather, the states are estimated through the observations of other information (e.g., default events, losses, etc.), as is often the case for the investors in the financial markets.

Our model differs from the previous contributions in the following sense, and has a number of advantages. Firstly, the default intensity  $\lambda$  in our model depends not only on the current state of the underlying state processes, but more importantly also on the previous default history of all the firms. With this approach, the contagion and frailty effects arise more naturally and explicitly. The interplay of the frailty and the default contagion effects is captured in the switching underlying state process, as well as the dependence of  $\lambda$  on the previous history. Secondly, the state process in our model is an unobservable process driving default intensities.<sup>1</sup> This represents a problem faced by investors in the market, namely, that they can observe only incomplete accounting information. We model the unobservable

---

<sup>1</sup>In Azizpour, Giesecke and Schwenkler (2014), an observable factor is also included in the model to represent the macro-economic risk factor. Our model and filtering-based approach can be easily extended to allow for observable risk factors, as observable factors will not change the filtering equations essentially. For simplicity, we only model the latent frailty and contagion effects through the hidden, unobservable Markov state process, as our main focus is to filter out the current state with the “self-exciting” default intensity, which is a novel contribution. We also obtain the robust filter with the “self-exciting” default intensity.

state processes as continuous-time, finite-state Markov chains.<sup>2</sup> We choose to use finite-state Markov chains due to its tractability, which enables us to obtain exact filters and robust filters of the current state. Meanwhile, it is known that a continuous-time Itô process can always be **approximated** by discrete-state Markov chains. Thirdly, with our filter and robust filter, we are able to obtain closed-form expressions for the price of credit risk bearing securities, with the “self-exciting” default intensity. Previous contributions are mostly concerned with default rates. Also, we consider parameter estimation using the EM-algorithm.<sup>3</sup> The EM algorithm is known to converge to the ML estimators reasonably rapidly.

## 3.2 Model Dynamics and Methodology

In this section, we describe our model dynamics, which allow for both *contagion* and *frailty* by using “self-exciting” process. We shall also briefly summarize our methodology and tools.

Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a complete, filtered probability space, with the filtration  $\mathbb{F} = (\mathcal{F})_{t \geq 0}$  satisfying the usual conditions. We suppose that the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  is rich enough to model the randomness of the observations process and of the state processes. Thus, all stochastic processes considered are, by definition,  $\mathbb{F}$ -adapted. We shall use this fact in various places in this paper.

The probability measure  $\mathbb{P}$  is assumed to be the historical (the “real world”, or physical) measure. The risk-neutral measure where the pricing of securities takes place is denoted by  $\mathbb{Q}$ . The measures  $\mathbb{P}$  and  $\mathbb{Q}$  are linked by the pricing kernel (state price density, or stochastic discount factor)

---

<sup>2</sup>In Azizpour, Giesecke and Schwenkler (2014), the frailty factor is modeled as a mean-reverting CIR square-root diffusion process.

<sup>3</sup>In Giesecke and Schwenkler (2014), the authors consider a filtered likelihood estimate of parameters for marked point processes. Using numerical analysis, they compare estimates obtained with their approach and with the EM algorithm, and find that for the majority of the parameters, their estimates are more accurate and less susceptible to the initial values.

### 3.2.1 Model Dynamics with Contagion and Frailty

Suppose we have  $N$  defaultable securities issued by  $N$  firms. We are concerned with the credit derivatives, i.e., derivative securities whose payoffs depends on the *default events* of the  $N$  firms.

Suppose there exists a state process  $(X_t)_{t \geq 0}$  that drives the common dynamics of the credit events (and, therefore, of the defaultable securities) of the  $N$  firms. In the context of our model, the states of processes  $X$  represent different economic environments, and firm specific characteristics. The switching of the states of the economy can be attributed to structural changes in macro-economic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities.

For tractability, we suppose that the process  $X := (X_t)_{t \geq 0}$  is a continuous-time, finite-state, Markov process, defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , with state space  $S$ . Without loss of generality, we take the state space of  $X$  to be the set  $\{e_1, \dots, e_k\} \subseteq \mathbb{R}^k$  of unit vectors, where  $e_i := (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^k$  with the “1” in the  $i$ th entry and where the superscript  $\top$  represents the transpose of a row vector. This is called the *canonical representation* of the state space of the Markov chain  $X$ .

Let  $A$  denote a constant rate matrix, or  $Q$ -matrix, of the Markov chain  $X$ . Then, using the canonical representation of the state space, the dynamics of the Markov chain  $X$  have the following *semi-martingale representation*

$$dX_t = AX_t dt + dM_t, \tag{3.1}$$

where  $(M_t)_{t \geq 0}$  is an  $\mathbb{R}^k$ -valued *martingale* with respect to the natural filtration generated by  $(X_t)_{t \geq 0}$ .

For  $1 \leq i \leq N$ , the random time  $\tau_i$  denotes the default time of firm  $i$ . The default event of firm  $i$  is assumed to depend on the intensity process  $(\lambda_t^i)_{t \geq 0}$ .

The *percentage loss*, given the default of firm  $i$ , will be denoted by the random variable

$\ell_i \in (0, \infty)$ . We assume  $\ell_1, \dots, \ell_N$  are independent random variables. The loss size  $\ell_i$  is a random variable with a density function  $f^i(\tau_i, \ell)$ . Here the parameter  $\tau$  in the function  $f^i(\tau, \ell)$  indicates that the density may depend on the default time  $\tau_i$ . For  $1 \leq i \leq N$ , define  $Z_t^i := \ell_i \mathbf{1}_{\{\tau_i \leq t\}}$ . Then the *loss state* of the credit derivatives is given by the multi-dimensional vector process

$$(Z_t)_{t \geq 0} = (Z_t^1, \dots, Z_t^N)^\top.$$

Let  $(N_t)_{t \geq 0}$  be a counting process which gives the number of firms that have defaulted by time  $t$ . We can write

$$N_t = \sum_{i=1}^N \mathbf{1}_{\{t \geq \tau_i\}}.$$

Similarly, write  $N_t^i = \mathbf{1}_{\{t \geq \tau_i\}}$  for the process which indicates whether firm  $i$  has defaulted by time  $t$ .

Suppose for each firm  $i$ , the default intensity of firm  $i$  at time  $t$ ,  $\lambda_t^i = \lambda^i(t, X_{t-}, N_{(\cdot)})$ , depends on  $X_{t-}$  and the previously realized process  $\{N_s : 0 \leq s < t\}$ . Note that only the state an instant ago  $X_{t-}$  appears in the intensity function, but the previously observed process  $N$  can affect the rate in a general way.<sup>4</sup>

We note that, the process  $X$  and the intensity  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  model the “*contagion*” and “*frailty*” effects, by which the default (or even the state switching) of one firm could have a direct influence on the default intensity of other firms, or on the other hand, many firms could be jointly exposed to one or more unobservable risk factors. In this sense, our model allows for both *contagion* and *frailty*.

The filtration  $\{\mathcal{F}_t^X, t \geq 0\}$  is defined by  $\mathcal{F}_t^X := \sigma\{X_s : s \leq t\}$ . Similarly, we write  $\mathcal{F}_t^Z = \sigma(Z_s; 0 \leq s \leq t)$ . Note that in our setting,  $\mathcal{F}_t = \sigma(X_s, Z_s; 0 \leq s \leq t)$ .

---

<sup>4</sup>For  $\lambda_t^i$  of the special form  $\lambda^i(t, X_{t-}, N_{(\cdot)}) = \phi\left(\int_{-\infty}^t h(t-s, X_{t-}) dN_s\right)$ , appropriate conditions for integrability are given by researchers. A key example of interest is when  $\lambda^i(t, X_{t-}, N_{(\cdot)}) = \langle \alpha, X_{t-} \rangle + \langle \beta, X_{t-} \rangle \int_0^t e^{-\langle \gamma, X_{t-s} \rangle (t-s)} dY_s^i$ , where  $Y_t^i = \sum_{j \neq i} N_t^j = \sum_{j \neq i} \mathbf{1}_{\{t \geq \tau_j\}}$ , and  $\alpha, \beta, \gamma$  are vectors of parameters to be determined. This is a variant of the Hawkes’ process.



### 3.2.2 Simulation of Default times

In our model, the default intensity is itself a stochastic process, and therefore, the simulation of the random default time  $\tau$  is worthy of consideration, from a computational perspective. The simulation of a Markov chain is routine. In this section, we briefly summarize two ways of simulating default times. For details, see Duffie and Singleton (2003). We adapt the following two well-known algorithms for simulation of the default time  $\tau$ .

The first method is known as the *inverse CDF simulation*. Let the survival probability of  $\tau$  be  $p(t) = \mathbb{P}(\tau > t)$ . We simulate a uniform distributed random variable  $\eta$  and let  $\tau$  be chosen so that

$$\tau := \inf \{t : \mathbb{P}(\tau > t) \geq \eta\}.$$

Alternatively, one can use what is known as *compensator simulation* method. Let  $q(t) := \int_0^t \lambda_u du$ . Simulate, independently of the intensity process  $\lambda$ , a standard exponentially distributed random variable  $\eta$  with unit-mean. We set

$$\tau := \inf \left\{ t : \int_0^t \lambda_u du \geq \eta \right\}.$$

For more detailed explanation of other approaches for the simulation of default times, the reader is referred to Duffie and Singleton (2003).

### 3.2.3 Point Processes and Random Measure

The pricing of credit derivatives in our model will be modeled by random measures. In this subsection, we give a brief discussion on point processes and random measures that will be used in this paper.

Recall that for each firm  $i$ ,  $1 \leq i \leq N$ , the default time is denoted  $\tau_i$ . Then each pair

$$(\tau_i, \ell_i), \quad 1 \leq i \leq N,$$

gives a representation of the single-jump process  $(Z_i^i)_{t \geq 0}$  as a *marked point process*, with the *mark space*  $E := (0, \infty)$ . We assume  $\ell_1, \dots, \ell_N$  are random variables, with a density function

$f^i(\tau, \ell)$ . Note that when the  $i$ th firm defaults at time  $\tau_i$ , the density function  $f^i(\tau_i, \ell)$  for the random loss size depends on the random time  $\tau_i$ . Also, the density function  $f^i(\tau_i, \ell)$  depends on the states  $(X_{\tau_i-}, X_{\tau_i-}^1, X_{\tau_i-}^2, \dots, X_{\tau_i-}^N)$ .

For each  $i$ ,  $1 \leq i \leq N$ , suppose  $\mu^i(\cdot, \cdot)$  is a *random measure* on  $[0, \infty) \times E$ , which gives the random time of the  $i$ th firm's default and the random loss size  $\ell_i \in (0, \infty)$ ,  $1 \leq i \leq N$ . More precisely, the random measure  $\mu^i(dt, de)$  is a random delta function

$$\mu^i(dt, de) = \delta_{\tau_i}(dt)\delta_{\ell_i}(de),$$

where  $\delta_x(\cdot)$  is a Dirac delta function, or a point mass, at the point  $x$ .

For a function  $g : \Omega \times [0, \infty) \times E \rightarrow \mathbb{R}$ ,

$$\int_0^t \int_E g(\omega, u, e) \mu^i(du, de) = g(\omega, \tau_i, \ell_i) \mathbf{1}_{\{\tau_i \leq t\}}.$$

Recall

$$Z_t^i = \mathbf{1}_{\{t \geq \tau_i\}} \ell_i = \begin{cases} 0 & t < \tau_i, \\ \ell_i & t \geq \tau_i. \end{cases}$$

Then the loss of firm  $i$  at time  $t$  can be written as

$$Z_t^i = \int_0^t \int_E \ell_i \mu^i(du, de) = \ell_i \mathbf{1}_{\{\tau_i \leq t\}}.$$

Hence, the cumulative loss of all the  $N$  firms, by time  $t$ , can be written as

$$\sum_{i=1}^N \int_0^t \int_E \ell_i \mu^i(du, de).$$

In particular, let  $N_t$  denote the number of default events up to time  $t$ . Then, in terms of the random measures  $\mu^i(\cdot, \cdot)$ , the counting process  $N_t$  is

$$N_t = \sum_{i=1}^N \int_0^t \int_E \mu^i(du, de).$$

### 3.2.4 Dividend Streams

In order for us to determine the fair price of a credit derivative, information about a *dividend stream* is needed.

Suppose there are  $N$  liquidly traded credit derivatives in the credit market. Most credit derivatives feature various kinds of intermediate cash flows,  $D_t^i$ , including payments at default dates, either to the holders or the issuers of the derivative.

For  $1 \leq i \leq N$ , we can describe the payoff of the derivatives of firm  $i$  by the cumulative dividend stream  $D_t^i$ . For many derivatives, the dividend stream  $D_t^i$  can be written, using the random measure, as

$$D_t^i = \int_0^t \phi_i(u, Z_u) dg(u) + \int_0^t \int_E \varphi_i(u, Z_{u-}, \ell) \mu^i(du, de).$$

Here  $\phi_i(\cdot, \cdot)$ ,  $\varphi_i(\cdot, \cdot, \cdot)$  are bounded functions on  $[0, \infty) \times [0, 1]^m$  and  $[0, \infty) \times [0, 1]^m \times [0, \infty)$ , respectively, and  $g(\cdot)$  is an increasing deterministic function  $g : [0, T] \rightarrow \mathbb{R}$ . We note that,  $\phi_i(\cdot, \cdot)$  and  $\varphi_i(\cdot, \cdot, \cdot)$  capture the cash flows that the holder of the credit derivative receives (or pays out) at the pre-scheduled regular times or when default occurs, respectively.

Dividend streams of this form can model many important credit derivatives, including bonds, credit default swaps, and collateralized debt obligations, to name just a few.

For example, for a zero-coupon bond with *zero recovery* paying  $y$  at the maturity date  $T$ , we have  $g(u) = \mathbf{1}_{\{u \geq T\}}$ ,  $\phi(u, Z_{u-}) = \mathbf{1}_{\{\tau \geq u\}}y = \mathbf{1}_{\{Z_u^i=0\}}y$ , and  $\varphi(u, Z_{u-}, \ell) = 0$ .

For a *credit default swap (CDS)* on a firm's bond, the holder of the CDS pays regular payments of size  $S$  to the issuer of the CDS at the pre-scheduled regular payment times  $t_1 < t_2 < \dots < t_m$  unless the firm defaults; in exchange, the issuer of the CDS pays the loss  $\ell$  to the holder if the firm defaults. Thus the dividend stream  $D_t$  of the CDS can be obtained if we take  $g(u) = |\{i : t_i \leq u\}|$  (where  $|\cdot|$  denotes the cardinality of a set) which gives the number of the payments by time  $t$ ,  $\phi(u, Z_u) = -\mathbf{1}_{\{\tau \geq u\}}S = -\mathbf{1}_{\{Z_u=0\}}S$  which models the regular payments at the pre-scheduled times  $t_1 < t_2 < \dots < t_m$ , and  $\varphi_i(u, Z_{u-}, \ell) = \mathbf{1}_{\{u \leq T\}}\mathbf{1}_{\{\tau=u\}}\ell$  which models the default loss if the firm defaults.

For the detailed examples of expressing a dividend stream, see Section 3.4.3. The reader is also referred to Bielecki and Rutkowski (2002), or Frey and Schmidt (2012).

### 3.2.5 Payoffs of Credit Derivatives

Consider a market of  $N$  credit derivatives with common maturity  $T$ . With *cumulative* dividend streams  $D_t^i$  given,  $1 \leq i \leq N$ , the discounted cumulative dividend streams are

$$\tilde{D}_t := \int_0^t e^{-\int_0^s r_u du} dD_s.$$

Then the discounted payoff of a credit derivative when it matures is (it is indeed the total sum of discounted payments in the time interval  $(t, T]$ )

$$H_{t,T}^i := \int_t^T d\tilde{D}_s = \tilde{D}_T^i - \tilde{D}_t^i = \int_t^T e^{-\int_0^s r_u du} dD_s.$$

The *intrinsic value* (or *hypothetical value*) of a credit derivative can be determined as follows. The filtration  $(\mathcal{F}_t)_{t \geq 0}$ ,  $\mathcal{F}_t := \sigma\{(X_s, Z_s) : s \leq t\}$ , is generated by the processes  $(X_t)_{t \geq 0}, (Z_t)_{t \geq 0}$ , up to time  $t$ . We are working with discounted quantities already. The **intrinsic value** (or **hypothetical value**) of a credit derivative with discounted payoff  $H_t$  is thus defined to be  $\mathbb{E}^{\mathbb{Q}}(H_t | \mathcal{F}_t)$ , the expectation conditional on the full information  $\mathcal{F}_t$  up to time  $t$ . We write  $p_t^i$  for the intrinsic value of the credit derivative, i.e.,

$$p_t^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} \left[ \tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t \right].$$

The computation of the full information value  $p^i$  can be done in a number of ways; see the remark after Proposition 3.4.1.

The market information is generated by the observation process

$$(Z_t)_{t \geq 0} = (Z_t^1, Z_t^2, \dots, Z_t^N).$$

The market information  $(\mathcal{F}_t^Z)_{t \geq 0}$  is given by  $\mathcal{F}_t^Z := \sigma\{Z_t : t \geq 0\}$ . Note that we assumed  $(\mathcal{F}_t)_{t \geq 0}$  is rich enough so that all stochastic processes are  $(\mathcal{F}_t)$ -adapted. Hence  $\mathcal{F}_t^Z \subseteq \mathcal{F}_t$ .

The **market value** of a credit derivative with discounted payoff  $H_{t,T}^i$  is thus  $\mathbb{E}[H_{t,T}^i|\mathcal{F}_t^Z]$ , the expectation conditional on the history  $\mathcal{F}_t^Z$  of the investors' observed information. We write  $\widehat{p}_t^i$  for the market value of a credit derivative, i.e.,

$$\widehat{p}_t^i := \mathbb{E}^{\mathbb{Q}}[H_{t,T}^i|\mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] = \mathbb{E}^{\mathbb{Q}} \left[ \widetilde{D}_T^i - \widetilde{D}_t^i | \mathcal{F}_t^Z \right].$$

Note that  $\widehat{p}_t^i$  is  $\mathcal{F}_t^Z$ -measurable (thus  $\widehat{p}^i$  is  $\mathbb{F}^Z$ -adapted). However,  $p^i$  is not  $\mathbb{F}^Z$ -adapted, as  $p_t^i$  is not necessarily  $\mathcal{F}_t^Z$ -measurable.

### 3.3 Estimation of Current State

#### 3.3.1 Filter Equation

Let  $\widetilde{\mathbb{P}}$  be a measure under which, for all  $1 \leq i \leq N$ ,  $Z^i$  is a standard single-jump process with unit intensity and jump size 1. Let  $\Lambda^i$  be the solution of the equation

$$\begin{aligned} \Lambda_t^i &= 1 + \int_0^t \Lambda_{u-}^i (\lambda_u^i - 1) (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du) \\ &= \exp \left( - \int_0^t (\lambda_u^i - 1) \mathbf{1}_{\{u \leq \tau_i\}} du \right) \left( \prod_{0 \leq u \leq t} \lambda_u^i \right)^{\Delta Z_u^i}. \end{aligned}$$

Define

$$\begin{aligned} \Lambda_t &= \prod_{i=1}^N \Lambda_t^i \\ &= \prod_{i=1}^N \exp \left( - \int_0^t (\lambda_u^i - 1) \mathbf{1}_{\{u \leq \tau_i\}} du \right) \left( \prod_{0 \leq u \leq t} \lambda_u^i \right)^{\Delta Z_u^i}. \end{aligned}$$

Then

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \Lambda_{u-} (\lambda_u^i - 1) (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du).$$

**Lemma 3.3.1** *Let  $\widetilde{\mathbb{P}}$  be a measure under which  $Z^i$ 's are all standard single-jump processes with unit intensity and unit jump size. Then under the measure  $\mathbb{P}$  defined by*

$$\left. \frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}} \right|_{\mathcal{F}_t} = \Lambda_t,$$

*$Z^i$  has intensity  $\lambda_u$ .*

*Proof.* See Appendix. □

We wish to determine a recursive relation for  $\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}^Z]$ . By a version of Bayes' Theorem (see Elliott, Aggoun and Moore (1995)), we have

$$\mathbb{E}^{\mathbb{P}}[X_t|\mathcal{F}^Z] = \frac{\mathbb{E}^{\tilde{\mathbb{P}}}[\Lambda_t X_t|\mathcal{F}^Z]}{\mathbb{E}^{\tilde{\mathbb{P}}}[\Lambda_t|\mathcal{F}^Z]}.$$

We shall find the unnormalized density

$$q_t = \mathbb{E}^{\tilde{\mathbb{P}}}[\Lambda_t X_t|\mathcal{F}^Z],$$

so that  $\mathbb{E}^{\mathbb{P}}[\Lambda_t|\mathcal{F}^Z] = \langle q_t, \mathbf{1} \rangle$ .

**Theorem 3.3.2** *The unnormalized density process  $q$  satisfies the  $\mathcal{F}^Z$ -adapted stochastic vector-valued ODE*

$$\begin{aligned} q_t &= q_0 + \sum_{i=0}^N \int_0^t \text{diag}(\lambda^{(i)} - 1) q_{u-} (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du) + \int_0^t A_u q_u du \\ &= q_0 + \sum_{i=0}^N \int_0^t \text{diag}(\lambda^{(i)} - 1) q_{u-} dZ_u^i - \sum_{i=0}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda^{(i)} - 1) q_{u-} du + \int_0^t A_u q_u du, \end{aligned}$$

where  $\text{diag}(\lambda^{(i)} - 1)$  is the diagonal matrix with  $(k, k)$ -entry  $\lambda_{ku}^i = \lambda^i(u, e_k, N_{(\cdot)})$ .

*Proof.* See Appendix. □

### 3.3.2 A Further Filter Equation

Suppose we have a reference probability measure  $\bar{\mathbb{P}}$  such that under  $\bar{\mathbb{P}}$ , for all  $1 \leq i \leq K$ ,  $Z_t^i$  is a marked point process with unit intensity for the random default times and a density function  $f(y)$ , independent of the hidden state  $X$ , for the loss sizes. The existence of such a reference probability measure  $\bar{\mathbb{P}}$  is shown in Elliott et al. (1995).

For each  $k = 1, 2, \dots, K$ , define

$$h_k^i(y) = \frac{\lambda_{k,u}^i f_k^i(y)}{f(y)} = \frac{\lambda^i(u, e_k, N_{(\cdot)}) f_k^i(y)}{f(y)}.$$

Define  $h_u^i(y) := \langle X_{u-}, (h_k^1(y), \dots, h_K^i(y)) \rangle$ . Thus  $h_u^i(y) = h_k^i(y)$  iff  $X_{u-} = e_k$ . Write  $H_k^i(u) := h_k^i(y_u)$ .

Then we have another filter equation, where we integrate against  $N_u^i$  rather than  $Z_u^i$ .

**Theorem 3.3.3** *Let  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^K$ . Suppose  $q_t$  is a  $K$ -dimensional vector such that*

$$\begin{aligned} q_t = & q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \text{diag}(H_1^i(u) - 1, \dots, H_K^i(u) - 1) q_u - dN_u^i \\ & - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda_1^i - 1, \dots, \lambda_K^i - 1) q_u du. \end{aligned}$$

Then

$$\mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z] = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}.$$

*Proof.* See Appendix. □

### 3.3.3 Robust Filters

In this section, we consider a “robust” filter, where we can avoid integrating against  $N^i$  or  $Z^i$  in the filtering equation.

We first introduce a gauge transformation matrix  $\Gamma_t$ . For  $k = 1, 2, \dots, K$ , define the scalar-valued process  $\gamma_k := \{\gamma_k(t)\}_{0 \leq t \leq T}$  by

$$\begin{aligned} \gamma_k(t) &:= \exp \left( \sum_{i=1}^N \int_0^{t \wedge \tau_i} (1 - \lambda^i(u, e_k, N_{(\cdot)})) du + \sum_{i=1}^N \int_0^t \log(H_{k,u}^i) dN_u^i \right) \\ &= \exp \left( \sum_{i=1}^N \int_0^{t \wedge \tau_i} (1 - \lambda_{k,u}^i) du + \sum_{i=1}^N \int_0^t \log(H_{k,u}^i) dN_u^i \right), \quad 0 \leq t \leq T \\ &= \exp \left( - \sum_{i=0}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda_{k,u}^{(i)} - 1) du \right) \prod_{i=1}^N \left( \prod_{0 \leq u \leq t} H_{k,u}^i \right)^{\Delta N_u^i}. \end{aligned}$$

where we write  $\lambda_{k,u}^i := \lambda^i(u, e_k, N_{(\cdot)})$  and

$$H_{k,u}^i = H_k^i(u) = \frac{\lambda_{k,u}^i f_k^i(y_u)}{f(y_u)}.$$

Define

$$\Gamma_t := \text{diag}(\gamma_1(t), \gamma_2(t), \dots, \gamma_K(t)).$$

Then, the inverse matrix  $\Gamma_t^{-1}$  of  $\Gamma_t$  is

$$\Gamma_t^{-1} := \text{diag}(\gamma_1^{-1}(t), \gamma_2^{-1}(t), \dots, \gamma_K^{-1}(t)).$$

**Lemma 3.3.4** For  $i = 1, 2, \dots, N$ , write

$$\begin{aligned} \text{diag}(\lambda^{(i)} - 1) &:= \text{diag}(\lambda_{1,u}^i - 1, \lambda_{2,u}^i - 1, \dots, \lambda_{K,u}^i - 1), \\ \text{diag}\left(\frac{1}{H^{(i)}} - 1\right) &:= \text{diag}\left(\frac{1}{H_{1,u}^i} - 1, \frac{1}{H_{2,u}^i} - 1, \dots, \frac{1}{H_{K,u}^i} - 1\right), \\ \text{or } \text{diag}\left(\frac{1 - H^{(i)}}{H^{(i)}}\right) &= \text{diag}\left(\frac{1 - H_{1,u}^i}{H_{1,u}^i}, \frac{1 - H_{2,u}^i}{H_{2,u}^i}, \dots, \frac{1 - H_{K,u}^i}{H_{K,u}^i}\right). \end{aligned}$$

Then we have

$$d(\Gamma_t^{-1}) = \sum_{i=1}^N \mathbf{1}_{\{t \leq \tau_i\}} \text{diag}(\lambda^{(i)} - 1) \Gamma_t^{-1} dt + \sum_{i=1}^N \text{diag}\left(\frac{1}{H^{(i)}} - 1\right) \Gamma_t^{-1} dN_t^i,$$

where  $\Gamma_0 = \Gamma_0^{-1} = I$ , the  $K \times K$  identity matrix.

*Proof.* See Appendix. □

Define the robust filter  $\bar{q}_t$  to be

$$\bar{q}_t := \Gamma_t^{-1} q_t.$$

Then we have the following recursive equation for the robust filter. Note that this equation does not directly involve the observation process  $Z$ , as it only appears through  $\Gamma$ .

**Theorem 3.3.5** The quantity  $\bar{q}_t$  satisfies the recursive equation

$$\bar{q}_t = \bar{q}_0 + \int_0^t \Gamma_u^{-1} A_u \Gamma_u \bar{q}_u du.$$



*Proof.* See Appendix. □

*Remark.* In the above theorem, we have an integral only against  $dt$ , and have avoided integrating against  $N^i$  or  $Z^i$  in the filtering equation. The above equation can be solved as an ODE, or using a routine numerical scheme, e.g., numerical integration, or the fast Fourier transform, which are widely available in various software packages. Once we have  $\bar{q}_t$ , it is relatively easy to obtain  $q_t$ .

## 3.4 Pricing of Defaultable Contingent Claims

### 3.4.1 Construction of General Dividends

In general, we consider a contingent claim (credit security) which has the following features. For a detailed description, the reader is referred to Bielecki and Rutkowski (2002).

The contingent claim has par value  $Y$ , which represents the payoff at maturity, if no default occurred prior to or at time  $T$ . If default occurred prior to or at time  $T$ , the holder of the claim receives a recovery claim  $Y'$  at time  $T$ . The holder of the security may also receive a recovery payoff  $R_\tau$  at the time of default  $\tau$ . The owner of the claim may as well receive (or pay out) intermediate, cumulative cash flow (e.g., bond yields, dividends, or CDS premium payouts), which is denoted by  $(C_t)_{0 \leq t \leq T}$ . Then the cumulative dividend process  $D$  of a defaultable contingent claim can be written as

$$\begin{aligned} D_t &= (Y\mathbf{1}_{\{\tau > T\}} + Y'\mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^t (1 - N_u) dC_u + \int_0^t R_u dN_u \\ &= (Y\mathbf{1}_{\{\tau > T\}} + Y'\mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^{t \wedge \tau} dC_u + R_\tau \mathbf{1}_{\{\tau \leq t\}}. \end{aligned}$$

### 3.4.2 Price of a Defaultable Contingent Claim

Consider a portfolio of  $N$  credit derivatives with common maturity  $T$ , with *cumulative* dividend streams  $D_t^i$  given.  $1 \leq i \leq N$ . The discounted cumulative dividend streams are

$$\tilde{D}_t^i := \int_0^t e^{-\int_0^s r_u du} dD_s^i.$$

The market value  $\widehat{p}_t^i$  of a credit derivative with discounted payoff  $H_{t,T}^i = \widetilde{D}_T^i - \widetilde{D}_t^i$  is the expectation conditional on the history  $\mathcal{F}_t^Z$  of the investors' observed information, i.e.,

$$\widehat{p}_t^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t^Z] = \mathbb{E}^{\mathbb{Q}} [\widetilde{D}_T^i - \widetilde{D}_t^i | \mathcal{F}_t^Z].$$

Also recall that

$$p_t^i := \mathbb{E}^{\mathbb{Q}} [H_{t,T}^i | \mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}} [\widetilde{D}_T^i - \widetilde{D}_t^i | \mathcal{F}_t^Z].$$

Note that  $D_T^i - D_t^i$  depends on the path  $\{Z_u : t < u \leq T\}$ . Also, by definition, the intrinsic value  $p_t^i$  depends on the state  $X_t$ . From the Markov property of the processes  $(X, Z)$ , we have (see, e.g., Chapter 2, Section 5 in Shreve et al. (1988))

$$\mathbb{E}^{\mathbb{Q}} [\widetilde{D}_T^i - \widetilde{D}_t^i | \mathcal{F}_t] = p^i(t, X_t, Z_t),$$

i.e.,  $p_t^i$  can be written as a function of  $X_t$  and  $Z_t$ .

The market price  $\widehat{p}_t^i$  of the credit derivative is given in the following proposition.

**Proposition 3.4.1** *The market value  $\widehat{p}_t^i$  of a credit derivative with dividend stream  $D_t^i$  is*

$$\begin{aligned} \widehat{p}_t^i &= \mathbb{E}^{\mathbb{Q}} [p^i(t, X_t, Z_t) | \mathcal{F}_t^Z] \\ &= (p_t^i(e_1, Z_t), p_t^i(e_2, Z_t), \dots, p_t^i(e_K, Z_t)) \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z]. \end{aligned}$$

*Proof.* Note that  $\mathbb{E}^{\mathbb{P}}(X_t | \mathcal{F}_t^Z)$  is a  $K$ -dimensional vector. By the definition of  $\widehat{p}_t^i$ , we have

$$\begin{aligned} \widehat{p}_t^i &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \Lambda_t \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{P}} \left[ \Lambda_t \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] && (\mathcal{F}_t^Z \subseteq \mathcal{F}_t) \\ &= \mathbb{E}^{\mathbb{P}} \left[ \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \middle| \mathcal{F}_t^Z \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^s r_u du} dD_s | \mathcal{F}_t \right] \mathbb{E}^{\mathbb{P}} [X | \mathcal{F}_t^Z] \\ &= \left( p_t^i(e_1, Z_t), p_t^i(e_2, Z_t), \dots, p_t^i(e_K, Z_t) \right) \cdot \mathbb{E}^{\mathbb{P}} [X_t | \mathcal{F}_t^Z]. \end{aligned}$$

Here, in the last equation, note that the product is the inner product of two  $K$ -dimensional vectors. □

### 3.4.3 Examples: Some Credit Securities

#### Defaultable Corporate Bonds

For a corporate defaultable bond, the cumulative dividend stream  $D_t$  can be modeled as follows. Suppose the bond has par value  $Y$ . If the firm defaults, then the recovery value of the bond at time  $T$  is  $Y' = \alpha Y$ . We denote the cumulative coupon by  $(C_t)_{0 \leq t \leq T}$ . Then the bond has cumulative dividend process

$$\begin{aligned} D_t &= (Y \mathbf{1}_{\{\tau > T\}} + Y' \mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^t (1 - N_u) dC_u \\ &= (Y \mathbf{1}_{\{\tau > T\}} + \alpha Y \mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_0^{t \wedge \tau} dC_u. \end{aligned}$$

The intrinsic price of the bond is

$$\begin{aligned} p^i(t, X_t, Z_t) &= \mathbb{E}^{\mathbb{Q}}[\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ (Y \mathbf{1}_{\{\tau > T\}} + Y' \mathbf{1}_{\{\tau \leq T\}}) \mathbf{1}_{\{t \geq T\}} + \int_t^T (1 - N_u) dC_u | \mathcal{F}_t \right]. \end{aligned}$$

The market price of the defaultable bond is

$$\hat{p}_t^i = \langle (p_t^i(e_1, Z_t), p_t^i(e_2, Z_t), \dots, p_t^i(e_K, Z_t)), \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z] \rangle.$$

#### Credit Default Swaps

A credit swap can be viewed as default insurance on reference loans or bonds. It pays the buyer of protection a given contingent amount at the time of a given credit event. If the insured event is a default, then the credit swap is known as a *credit default swap* (CDS). The contingent amount to be paid to the holder of a CDS at the time of the credit event (default, say) is referred to as *default payment*. In compensation for the default payment, the CDS holder pays the issuer of the CDS an annuity, at a rate variously called the credit-swap *spread*, CDS rate, or the credit-swap *premium*. This annuity stream is paid until the maturity of the CDS, or until the time of the credit event.

Suppose the protection buyer pays the annuity (the credit-swap premium) at times  $t_i$ ,  $i = 1, 2, \dots, m$ , prior to default or maturity, whichever is earlier. For simplicity, we postulate that reference asset is a zero-coupon bond with par value  $L$  and maturity  $T$ . The default payment at time  $t$  is then

$$(L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}\mathbf{1}_{\{\tau\}}(t) = (L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}\mathbf{1}_{\{t=\tau\}},$$

where  $D(\tau)$  is the recovery value of the reference bond at the time default. On the other hand, the premium payment at time  $t$  is

$$\sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{t_i\}}(t) = \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{t=t_i\}},$$

where  $\kappa$  denotes the annuity amount. Thus, from the CDS holder's perspective, the cumulative dividend cash flow by time  $t$  is

$$\begin{aligned} D_t &= \int_0^t (L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}\mathbf{1}_{\{\tau\}}(u)du - \int_0^t \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{t_i\}}(u)du \\ &= \int_0^t (L - D(\tau))\mathbf{1}_{\{\tau \leq T\}}dN_u - \int_0^t \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{u=t_i\}}du. \end{aligned}$$

When we postulate the fractional recovery  $\delta$  of par value, the cash flow can be further simplified as

$$D_t = \int_0^t L(1 - \delta)\mathbf{1}_{\{\tau \leq T\}}dN_u - \int_0^t \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{u=t_i\}}du.$$

The intrinsic price of the credit default swap is

$$\begin{aligned} p^i(t, X_t, Z_t) &= \mathbb{E}^{\mathbb{Q}}[\tilde{D}_T^i - \tilde{D}_t^i | \mathcal{F}_t] \\ &= \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_0^u r_s ds} L(1 - \delta)\mathbf{1}_{\{\tau \leq T\}}dN_u - \int_t^T e^{-\int_0^u r_s ds} \sum_{i=1}^m \kappa \mathbf{1}_{\{t_i < \tau\}} \mathbf{1}_{\{u=t_i\}}du \middle| \mathcal{F}_t \right]. \end{aligned}$$

The market price of the credit default swap is

$$\hat{p}_t^i = \langle (p_t^i(e_1, Z_t), p_t^i(e_2, Z_t), \dots, p_t^i(e_K, Z_t)), \mathbb{E}^{\mathbb{P}}[X_t | \mathcal{F}_t^Z] \rangle.$$

## Default Put Options

A default put option is similar to a credit default swap. The difference between a default put option and a CDS is the the protection buyer pays a lump sum *premium* upfront instead of an annuity to the issuer at the contract's inception. Therefore the cumulative cash flow by time  $t$  is

$$D_t = \int_0^t (L - D(\tau)) \mathbf{1}_{\{\tau \leq T\}} dN_u - \kappa.$$

The market price of a default option can be obtained similarly to that of a CDS.

## Other Credit Derivatives

Similarly we can derive the market price of other credit derivatives. Once the cumulative dividend stream is written out, it is relatively straightforward to obtain the market price, within our pricing framework. Other similar credit derivatives include, for example, credit linked notes (CLN), credit spread swaps, and credit spread options.

More complicated portfolio credit derivatives can be similarly priced, using the pricing equation in our model. For example, basket credit swaps, *i*th-to-default CDS, *m*-of-*n*-to-default CDS, synthetic CDO tranche, etc. The cumulative cash flow can be written out similarly to the pervious examples, with more involved analyses. The reader is referred to Bielecki and Rutkowski (2002) for more details about the cash flows for these basket credit derivative. Once the cumulative dividend stream is determined, the market price can be computed as before.

## 3.5 Parameter Estimation

We separate our estimation into two parts: estimations of the parameters for  $\lambda_i$ , and estimation of the rate matrix  $A$ . The reason is because we shall usually have a large number of observed jumps in  $Z$  relative to jumps in the underlying chain  $X$ . This implies that we

shall often be able to estimate the parameters for  $\lambda$  quite well given  $A$ . The estimation of  $A$ , which depends on the jumps of  $X$ , is usually much more costly.

Given the state path  $X$ , we proceed to find a maximum likelihood estimator, which can be combined with the EM algorithm. The standard log-likelihood functions given  $X$  and observations on the interval  $[0, T]$  is

$$\sum_{i=1}^N \sum_{u: \Delta Z_u^i \neq 0} \log(\lambda^i(t, X_{t-}, N_{(\cdot)})) - \int_0^T \lambda^i(t, X_{t-}, N_{(\cdot)}) du.$$

We can then get the partial-information likelihood function, given the probabilities  $r_u := \mathbb{E}[X_u | \mathcal{F}_u^Z]$ . Write  $\boldsymbol{\lambda}_u^i$  for the vector with entries  $\lambda^i(t, X_{t-}, N_{(\cdot)})$ .

**Lemma 3.5.1** *The log-likelihood function given observations on the time period  $[0, T]$  is*

$$\sum_{i=1}^N \sum_{u: \Delta Z_u^i \neq 0} \langle r_u, \log(\boldsymbol{\lambda}_u^i) \rangle - \int_0^T \langle r_u, \log(\boldsymbol{\lambda}_u^i) \rangle du.$$

This function can be optimized analytically, or numerically. One can use the EM algorithm and iterate to estimate the parameters for  $\lambda$  and calculate  $r$ . In each iteration, the estimation of the parameters for  $\lambda$  can be done using numerical optimization techniques, and calculating  $r$  given the parameters of  $\lambda$  can be done using the robust filtering equations.

To implement the estimation of  $A$  efficiently, we regard the rate matrix  $A$  as a tuning parameter for the filter. For example, we can choose a parameterized family of rate matrices, say

$$\left\{ \epsilon \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}; \epsilon > 0 \right\}.$$

Then, during the calibration, one chooses the value of  $\epsilon$  which gives reasonable performance of the filter in determining the underlying state. The effectiveness of this choice  $A$  could then be determined by out-of-sample performance.

On the other hand, if we need to also perform pricing of credit securities, then in addition to the parameters in  $\mathbb{P}$ , we would also need to know the parameters in the risk-neutral

measure  $\mathbb{Q}$ . The estimation of the parameters in the risk-neutral measure  $\mathbb{Q}$  can be performed by a parametric specification of the pricing kernel, which links the historical measure  $\mathbb{P}$  and the pricing measure  $\mathbb{Q}$ .

In general, we know that the pricing kernel  $M_t = e^{-\int_0^t r_u du} \Lambda_t$ , where  $\Lambda_t = d\mathbb{Q}/d\mathbb{P}$  is the Radon-Nikodym derivative, i.e.,

$$\Lambda_t = \exp \left( - \sum_{i=1}^N \int_0^t \int_E (h_u^i(y) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f_i^{\mathbb{P}}(y) du dy + \sum_{i=1}^N \int_0^t \int_E \log h_u^i(y) \mu^i(du, dy) \right).$$

Here

$$h_u^i(y) := \frac{\lambda_{iu}^{\mathbb{Q}} f_i^{\mathbb{Q}}(y)}{\lambda_{iu}^{\mathbb{P}} f_i^{\mathbb{P}}(y)} = \sum_{k=1}^K \langle X_{u-}, e_k \rangle \frac{\lambda_{ik}^{\mathbb{Q}} f_i^{\mathbb{Q}}(y)}{\lambda_{ik}^{\mathbb{P}} f_i^{\mathbb{P}}(y)} = \sum_{k=1}^K \langle X_{u-}, e_k \rangle h_k^i(y).$$

From the time series data of the market price of the securities, we obtain the parameters of the pricing kernel, and therefore obtain the parameters in the pricing measure  $\mathbb{Q}$ , with which we can calculate the securities' prices using the pricing measure  $\mathbb{Q}$ .

with this framework, we obtain joint estimation for the parameters uses time-series data of both historical defaults and security prices.

### 3.6 Concluding Remarks

In this paper, we investigated a credit risk model with self-exciting default intensity. Our models naturally allow for both frailty and default contagion.

We obtain the estimates of  $X_t$  and the pricing formulas for the credit contingent claims when the investors face incomplete market accounting information

Our method in the paper is very flexible. It can be *further generalized* to estimate  $X$  even if there are other noise observations, in addition to the loss state process  $Z$ . For example, our method can be applied if, in addition to  $Z$ , there are also observations consisting of functions of  $X$  and additive Gaussian noise.

We also investigate parameter estimation for the model. The parameter estimation is done using the EM (expectation maximization) algorithm.

## Appendix 1

*Proof of Theorem 3.3.2.* We have

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \Lambda_{u-}^i (\lambda_u^i - 1) (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du).$$

Then

$$\Lambda_t X_t = X_0 + \sum_{i=0}^N \int_0^t \Lambda_{u-} X_{u-} (\lambda_u^i - 1) (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du) + \int_0^t \Lambda_{u-} A_u X_u du + \int_0^t \Lambda_{u-} dM_u.$$

As  $\Lambda$  has independent increments under  $\tilde{\mathbb{P}}$ , taking a conditional expectation,

$$q_t = q_0 + \sum_{i=0}^N \int_0^t \tilde{\mathbb{E}}[\Lambda_{u-} X_{u-} (\lambda_u^i - 1) | \mathcal{F}_u^Z] (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du) + \int_0^t A_u q_u du.$$

As  $\lambda^i$  is a scalar function of the state  $X_{u-}$ , we can write

$$X_{u-} (\lambda_u^i - 1) = (\lambda^i(u, X_{u-}, N_{(\cdot)}) - 1) X_{u-} = \text{diag}(\lambda^{(i)} - 1),$$

where  $\text{diag}(\lambda^{(i)} - 1)$  is the diagonal matrix with  $(k, k)$ -entry  $\lambda_{k,u}^i = \lambda^i(u, e_k, N_{(\cdot)})$ . This implies that  $\text{diag}(\lambda^{(i)} - 1)$  is  $\mathcal{F}_u^Z$ -measurable. Substituting this in the equation for  $q$ , we have

$$\begin{aligned} q_t &= q_0 + \sum_{i=0}^N \int_0^t \text{diag}(\lambda^{(i)} - 1) q_{u-} (dZ_u^i - \mathbf{1}_{\{u \leq \tau_i\}} du) + \int_0^t A_u q_u du \\ &= q_0 + \sum_{i=0}^N \int_0^t \text{diag}(\lambda^{(i)} - 1) q_{u-} dZ_u^i - \sum_{i=0}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda^{(i)} - 1) q_{u-} du + \int_0^t A_u q_u du. \end{aligned}$$

□

*Proof of Theorem 3.3.3.* The proof is similar to that of Theorem 3.3.2. We start with  $\bar{\mathbb{P}}$ .

Using Girsanov's Theorem for jump processes, define a process  $\Lambda_t$  by

$$\begin{aligned} \Lambda_t &= \prod_{i=1}^N \exp \left( - \int_0^t \int_E X_{u-} (h_u^i(y) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y) du dy + \int_0^t \int_E X_{u-} \log h_u^i(y) \mu^i(du, dy) \right) \\ &= \exp \left( - \sum_{i=1}^N \int_0^t \int_E X_{u-} (h_u^i(y) - 1) \mathbf{1}_{\{u \leq \tau_i\}} f(y) du dy + \sum_{i=1}^N \int_0^t \int_E X_{u-} \log h_u^i(y) \mu^i(du, dy) \right). \end{aligned} \tag{3.2}$$



Using Itô's rule, we also have

$$\Lambda_t = 1 + \sum_{i=1}^N \int_0^t \int_E \Lambda_{u-} X_{u-} (h_u^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du),$$

or

$$d(\Lambda_t) = \Lambda_{t-} \sum_{i=1}^N \int_E X_{t-} (h_u^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy dt).$$

Since

$$\begin{aligned} d(\Lambda_t X_t) &= d(\Lambda_t) X_t + \Lambda_t d(X_t) \\ &= \Lambda_{t-} \left( \sum_{i=1}^N \int_E X_{t-} (h_u^i(y) - 1) (\mu^i(dy, dt) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy dt) \right) X_{t-} + \Lambda_{t-} (AX_t + dM_t), \end{aligned}$$

we have

$$\begin{aligned} \Lambda_t X_t &= \Lambda_0 X_0 + \sum_{i=1}^N \int_0^t \Lambda_{u-} X_{u-} \int_E X_{u-} (h_u^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) \\ &\quad + \int_0^t \Lambda_u A X_u du + \int_0^t \Lambda_{u-} dM_u. \end{aligned}$$

Write  $q_t := \overline{\mathbb{E}}[\Lambda_t X_t | \mathcal{F}_t^M]$ . Then, a version of Bayes' Theorem gives

$$\mathbb{E}[X_t | \mathcal{F}_t^M] = \frac{\overline{\mathbb{E}}[\Lambda_t X_t | \mathcal{F}_t^M]}{\overline{\mathbb{E}}[\Lambda_t | \mathcal{F}_t^M]} = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}.$$

Conditioning on  $\mathcal{F}_t^Z$  under  $\overline{\mathbb{P}}$ , using the Fubini Theorem in Hajek and Wong (1985), we have

$$q_t = q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle q_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) e_k$$

Write  $H_k^i(u) := h_k^i(y_u)$ . Then

$$\begin{aligned} q_t &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \int_E \sum_{k=1}^K \langle q_{u-}, e_k \rangle (h_k^i(y) - 1) (\mu^i(dy, du) - \mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) e_k \\ &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \int_E \text{diag}(h_1^i(y) - 1, \dots, h_K^i(y) - 1) q_{u-} \mu^i(dy, du) \\ &\quad - \sum_{i=1}^N \int_0^t \int_E \text{diag}(h_1^i(y) - 1, \dots, h_K^i(y) - 1) q_{u-} (\mathbf{1}_{\{u \leq \tau_i\}} f(y) dy du) \\ &= q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \text{diag}(H_1^i(u) - 1, \dots, H_K^i(u) - 1) q_{u-} \times \int_E \mu^i(dy, du) \end{aligned}$$

$$\begin{aligned}
& - \sum_{i=1}^N \int_0^t \text{diag}(\lambda_1^i - 1, \dots, \lambda_K^i - 1) \mathbf{1}_{\{u \leq \tau_i\}} q_u du \\
& = q_0 + \int_0^t A q_u du + \sum_{i=1}^N \int_0^t \text{diag}(H_1^i(u) - 1, \dots, H_K^i(u) - 1) q_u - dN_u^i \\
& \quad - \sum_{i=1}^N \int_0^{t \wedge \tau_i} \text{diag}(\lambda_1^i - 1, \dots, \lambda_K^i - 1) q_u du.
\end{aligned}$$

□

*Proof of Lemma 3.3.4.* By Itô's product rule,

$$\begin{aligned}
d(\gamma_k^{-1}(t)) & = d(\gamma_k^{-1}(t)) = -\gamma_k^{-1}(t-) \sum_{i=1}^N \left( (1 - \lambda_{k,t}^i) \mathbf{1}_{\{t \leq \tau_i\}} dt + \log(H_{k,t}^i) \Delta N_t^i \right) \\
& \quad + \gamma_k^{-1}(t-) \sum_{i=1}^N \left( \left( \frac{1}{H_{k,t}^i} - 1 \right) \Delta N_t^i + \log(H_{k,t}^i) \Delta N_t^i \right) \\
& = -\gamma_k^{-1}(t-) \sum_{i=1}^N \left( (1 - \lambda_{k,t}^i) \mathbf{1}_{\{t \leq \tau_i\}} dt + \log(H_{k,t}^i) \Delta N_t^i \right) \\
& \quad + \gamma_k^{-1}(t-) \sum_{i=1}^N \left( \left( \frac{1 - H_{k,t}^i}{H_{k,t}^i} \right) \Delta N_t^i + \log(H_{k,t}^i) \Delta N_t^i \right) \\
& = \gamma_k^{-1}(t-) \sum_{i=1}^N (1 - \lambda_{k,t}^i) \mathbf{1}_{\{t \leq \tau_i\}} dt + \gamma_k^{-1}(t-) \sum_{i=1}^N \left( \frac{1 - H_{k,t}^i}{H_{k,t}^i} \right) dN_t^i.
\end{aligned}$$

Thus

$$d(\Gamma_t^{-1}) = \sum_{i=1}^N \mathbf{1}_{\{t \leq \tau_i\}} \text{diag}(\lambda^{(i)} - 1) \Gamma_t^{-1} dt + \sum_{i=1}^N \text{diag} \left( \frac{1 - H^{(i)}}{H^{(i)}} \right) \Gamma_{t-}^{-1} dN_t^i.$$

□

*Proof of Theorem 3.3.5.* By Itô's product rule,

$$\begin{aligned}
d(\bar{q}_t) & = d(\Gamma_t^{-1} q_t) \\
& = \Gamma_t^{-1} d(q_t) + d(\Gamma_t^{-1}) q_t + d[\Gamma^{-1}, q]_t \\
& = \Gamma_t^{-1} \left( A_t q_t dt + \sum_{i=0}^N \text{diag}(H_1^i(t) - 1, \dots, H_K^i(t) - 1) q_t - dN_t^i - \sum_{i=0}^N \mathbf{1}_{t \wedge \tau_i} \text{diag}(\lambda^{(i)} - 1) q_t - dt \right) \\
& \quad + \left( \sum_{i=1}^N \mathbf{1}_{\{t \leq \tau_i\}} \text{diag}(\lambda^{(i)} - 1) \Gamma_t^{-1} dt + \sum_{i=1}^N \text{diag} \left( \frac{1}{H^{(i)}} - 1 \right) \Gamma_{t-}^{-1} dN_t^i \right) q_t
\end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^N \text{diag} \left( \frac{1}{H^{(i)}} - 1 \right) \Gamma_{t-}^{-1} \text{diag}(H^{(i)} - 1) q_{t-} dN_t^i \\
& = \Gamma_t^{-1} A_t q_t dt + \sum_{i=1}^N \left( \text{diag}(H^{(i)} - 1) + \text{diag} \left( \frac{1}{H^{(i)}} - 1 \right) + \text{diag} \left( \frac{1}{H^{(i)}} - 1 \right) \text{diag}(H^{(i)} - 1) \right) \Gamma_t^{-1} q_{t-} dN_t^i \\
& = \Gamma_t^{-1} A_t q_t dt.
\end{aligned}$$

Thus

$$\begin{aligned}
\bar{q}_t & = \bar{q}_0 + \int_0^t \Gamma_u^{-1} A_u q_u du \\
& = \bar{q}_0 + \int_0^t \Gamma_u^{-1} A_u \Gamma_u \bar{q}_u du.
\end{aligned}$$

□

## Appendix 2

In this appendix, we give a closed-form expression for the survival probability in our model with self-exciting intensities. We show that, if we exclude some uninteresting technical cases, a closed-form expression for the survival probability can be obtained. In our model,  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  depends on  $X_{t-}$ , and  $N_{[0, t-]}^j$ , for all  $j \neq i$ . We further assume the following, to exclude some uninteresting cases.

**Assumption.**  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  does not depend on  $N_{[0, t-]}^i$ .<sup>5</sup>

We make this assumption as we use  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  to model the contagion and frailty effects from other firms, so  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  does not depend on its own default history. Also, when  $N_{t-}^i = 0$ ,  $N_{[0, t-]}^i$  does not influence  $\lambda^i(t, X_{t-}, N_{(\cdot)})$ , and once  $N_{\tau_i}^i = 1$  for  $\tau_i < t$ , firm  $i$  has already defaulted. Thus, it is sensible to assume that  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  does not depend on  $N_{[0, t-]}^i$ .

---

<sup>5</sup>For example, a variant of the Hawkes' process is  $\lambda^i(t, X_{t-}, N_{(\cdot)}) = \langle \alpha, X_{t-} \rangle + \langle \beta, X_{t-} \rangle \int_0^t e^{-\langle \gamma, X_{t-s} \rangle (t-s)} dY_s^i$ , where  $Y_t^i = \sum_{j \neq i} N_t^j = \sum_{j \neq i} \mathbf{1}_{\{t \geq \tau_j\}}$ , and  $\alpha, \beta, \gamma$  are vectors of parameters to be determined. This is a parametric specification where  $\lambda^i(t, X_{t-}, N_{(\cdot)})$  is **self-exciting**, and not depending on  $N_{[0, t-]}^i$ .

With this assumption, we now prove the survival probability. By definition of stochastic intensity (see, e.g., Duffie and Singleton (2003) p. 361, Bielecki and Rutkowski (2002) p. 155), the compensator of  $N_t^i$  is  $\int_0^{t \wedge \tau} \lambda^i(u, X_{u-}, N_{(\cdot)}) du$ , i.e.,

$$N_t^i - \int_0^{t \wedge \tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \quad (3.3)$$

is an  $\mathcal{F}$ -martingale under measure  $\mathbb{P}$ . We need to find the survival probability  $\mathbb{P}(\tau_i \geq t | \mathcal{F}_t)$ .

The filtration  $\{\mathcal{F}_t^X, t \geq 0\}$  is defined by  $\mathcal{F}_t^X := \sigma\{X_s : s \leq t\}$ . Similarly, we write  $\mathcal{F}_t^Z = \sigma\{Z_s; 0 \leq s \leq t\}$ . Note that in our setting,  $\mathcal{F}_t = \sigma(X_s, Z_s; 0 \leq s \leq t)$ .

We have

$$\begin{aligned} \mathbb{E}[N_t^i | \mathcal{F}_s] &= \mathbb{E}\left[ I_{\{t \geq \tau_i\}} \middle| \mathcal{F}_s \right] \\ &= \mathbb{E}\left[ \int_0^{t \wedge \tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \middle| \mathcal{F}_s \right] \quad (\text{by (1)}) \\ &= \mathbb{E}\left[ I_{\{t \leq \tau_i\}} \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du + I_{\{t \geq \tau_i\}} \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \middle| \mathcal{F}_s \right] \\ &= \mathbb{E}\left[ I_{\{t \leq \tau_i\}} \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du \middle| \mathcal{F}_s \right] + \mathbb{E}\left[ I_{\{t \geq \tau_i\}} \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \middle| \mathcal{F}_s \right]. \end{aligned} \quad (3.4)$$

Now we calculate the first conditional expectation in equation (3.4). Since  $\lambda^i(u, X_{u-}, N_{(\cdot)})$  does not depend on  $N^i$  (or  $\tau_i$ ) conditional on  $\tau_i \geq t$ , we have

$$\begin{aligned} \mathbb{E}\left[ I_{\{t \leq \tau_i\}} \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du \middle| \mathcal{F}_s \right] &= \int_t^\infty I_{\{t \leq \tau_i\}} \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du d\mathbb{P}(\tau_i | \mathcal{F}_s) \\ &= \int_t^\infty \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du d\mathbb{P}(\tau_i | \mathcal{F}_s) \\ &= \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du \int_t^\infty d\mathbb{P}(\tau_i | \mathcal{F}_s) \\ &= \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du \left( 1 - \mathbb{P}(\tau_i < t | \mathcal{F}_s) \right). \end{aligned}$$

Now we calculate the conditional expectation in the second term in equation (3.4). For  $\tau_i \leq t$ ,  $\int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du$  depends on  $X_{u-}$ ,  $N_{[0, u]}^j$ ,  $j \neq i$ , and  $\tau_i$ . Then

$$\mathbb{E}\left[ I_{\{t \geq \tau_i\}} \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du \middle| \mathcal{F}_s \right] = \int_0^t I_{\{t \geq \tau_i\}} \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du d\mathbb{P}(\tau_i | \mathcal{F}_s)$$

$$= \int_0^t \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du d\mathbb{P}(s < \tau_i | \mathcal{F}_s).$$

Thus from equation (3.4) we have

$$\mathbb{P}(t | \mathcal{F}_s) = \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du \left(1 - \mathbb{P}(t | \mathcal{F}_s)\right) + \int_0^t \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du d\mathbb{P}(\tau_i | \mathcal{F}_s) \quad (3.5)$$

Denoting  $p(t) = \mathbb{P}(t | \mathcal{F}_s) = \mathbb{P}(s < t | \mathcal{F}_s)$ , we have from equation (3.5)

$$p(t) = \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du \left(1 - p(t)\right) + \int_0^t \int_0^{\tau_i} \lambda^i(u, X_{u-}, N_{(\cdot)}) du dp(\tau_i).$$

Differentiating the above equation, we have

$$\begin{aligned} dp(t) &= \lambda_t^i dt \left(1 - p(t)\right) - \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du dp(t) + \int_0^t \lambda^i(u, X_{u-}, N_{(\cdot)}) du dp(t) \\ &= \lambda_t^i dt \left(1 - p(t)\right). \end{aligned}$$

The above differential equation is equivalent to

$$d\left(1 - p(t)\right) = -\lambda_t^i dt \left(1 - p(t)\right),$$

which yields

$$1 - p(t) = \exp\left(-\int_0^t \lambda_u^i du\right).$$

Thus

$$\mathbb{P}(\tau_i > t | \mathcal{F}_s) = 1 - p(t) = \exp\left(-\int_0^t \lambda_u^i du\right).$$

This proves the desired probability. □

## Chapter 4

# Dynamic Optimal Capital Structure with Regime Switching

### 4.1 Introduction

Capital structure theory is among the central topics of finance theory. The value of corporate debt (and therefore, the credit spread) is closely connected to the firm's capital structure. Debt values cannot be determined without the knowledge of the firm's capital structure, which also affects default decision and bankruptcy. On the other hand, capital structure can only be optimized when the effect of leverage on the value of debt is known.

Traditional capital structure theory, initiated by Modigliani and Miller (1958), identified some prime determinants of optimal capital structure. This theory, however, has been less useful in practice as it provides only qualitative insights. Brennan and Schwartz (1978) gave the first quantitative investigation of optimal leverage. Their approach relies heavily on numerical techniques to determine optimal leverage when a firm's unlevered value follows a diffusion process with constant volatility.

Leland (1994), Leland and Toft (1996) made breakthrough in this direction, adopting the same assumptions for modeling as in Modigliani and Miller (1958), Merton (1974), Black and Cox (1976), and Brennan and Schwartz (1978). Leland (1994) considers endogenous bankruptcy which is triggered by the firm failing to fulfil its current debt obligation, and derives a closed-form solution for the time-independent corporate debt, e.g., a consol bond. In Leland and Toft (1996), the authors further examine corporate debts with finite expiration. Leland (1998) then considers strategic restructuring/renegotiation for the firm's debt.

The models considered by Merton (1974), Black and Cox (1976), Brennan and Schwartz

(1978), Leland (1994), Leland and Toft (1996), etc., model the evolution of firm value as a diffusion process. This assumption results in unrealistic empirical implications. For example, if a firm cannot default unexpectedly and if it is not currently in financial distress, its probability of defaulting on very short-term debt is zero and therefore, its short-term debt should have zero credit spreads and its term structure of credit spreads should slope upward at the short end. Empirical evidence indicates that credit spreads on typical short term bonds are essentially greater than zero. Moreover, Fons (1994) and Sarig and Warga (1989) show that the yield spread curves of certain kinds of bonds are even downward sloping. The empirical analysis in Jones, Mason, and Rosenfeld (1984) indicate that the credit spreads on corporate bonds are generally too high to be matched by this approach. Other empirical contributions that actually implement structural models include Delianedis and Geske (1999), Eom et al. (2004), Ericsson and Reneby (2002, 2005) and Huang and Huang (2012).

Researchers have proposed extensions of the approach employed in Merton (1974), Black and Cox (1976), Brennan and Schwartz (1978), etc. For example, the approach is later extended to allow for stochastic interest rates in Longstaff and Schwartz (1995), Kim et al. (1993), Cathart and El-Jahel (1998), Briys and de Varenne (1997), Nielsen et al. (1993), Saá-Requejo and Santa-Calra (1999), etc.; in Zhou (2001), the author models the evolution of the firm value as a jump-diffusion process, as introduced in Merton (1973).

On the other hand, there have been a large amounts of empirical evidence suggesting that the aggregate economy is characterized by periodic shifts between distinct business cycles (see, e.g., Hamilton (1988)). In addition to the empirical evidence, there are also economic reasons to believe that regime switching is important to understand the dynamics of the credit spread curve. Empirically, the expansion and contraction periods in business cycles have potentially sizable effects on macroeconomic fundamentals, such as inflationary expectations, monetary policy, and nominal interest rates. Also, these regime shifts change corporations' growth prospects, and affect corporate earnings, as well as investment or default

decisions. On economic grounds, such switchings in regimes give rise to the possibility of significant impacts on firms' profitability, riskiness of assets and investments, and the likelihood of debt defaults. Therefore, the regime switchings have impact on firms' capital accumulation, and policy choices, as well as investment or default decisions, and consequently affect the credit risk associated with these firms, as well as the credit spreads of the credit risk bearing securities issued by the firms. In particular, Giesecke et al. (2011) conducted empirical analyses on the corporate bond market over the past 150 years. Their empirical data suggests that there exist three regimes, and they also obtain the transition probabilities.

In a typical capital structure model, e.g., Merton (1974) and Black and Cox (1976), Brennan and Schwartz (1978), Leland (1994) and Leland and Toft (1996), etc., debt/equity leverage is analyzed by viewing debts and equities as contingent claims written on real assets. Because the values of contingent claims depend on the riskiness of the underlying assets, volatility is an important determinant of the default risk. In previous contributions, e.g., Merton (1974), Black and Cox (1976), Brennan and Schwartz (1978), Leland (1994), Leland and Toft (1996), Longstaff and Schwartz (1995), etc., models typically postulate that this parameter is a fixed constant. It is natural to imagine that, as interpreted above, as volatility changes over the business cycle, so does the capital structure and bankruptcy decision. Indeed, Huang and Zhou (2008) conducted empirical analyses to test these models, and their empirical test strongly reject the models.

In this paper, we develop a framework to model the capital structure when the dynamics of the state variables are subject to regime shifts at random times. Following Hamilton (1989), we think of these regime switchings as "episodes across which the behavior of the series is markedly different". We construct two models that allow for the regime switching risk, which extend earlier contributions in Black and Cox (1976), Brennan and Schwartz (1978), Leland (1994), etc. Moreover, we are able to obtain closed-form explicit solutions for the optimal capital structure in these models, where regime switchings are allowed.



The paper is organized as follows. Section 2 presents a basic model where tax effects and bankruptcy costs are not considered. This is a model extending the analysis in Black and Cox (1976). In section 3, we consider a model with endogenous default which takes tax effects and bankruptcy costs into account, and regime switching risk is also considered. This model extends the classic work in Leland (1994). For both models, we are able to obtain closed-form analytical solutions for the optimal capital structure. In section 4 we present concluding remarks.

## 4.2 Asset Dynamics and Coupled PDE under Regime Switching

### 4.2.1 Asset Dynamics under Regime Switching

Suppose the state of the economy is described by a finite state continuous-time Markov chain  $X = \{X_t, t \geq 0\}$ . In our models, the states of the Markov Chains represent different economic environments. For example, there could be just two states for  $X$  representing “good” and “bad” economic regimes. The switching of the states of the economy can be attributed to structural changes in macroeconomic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities.

For notational convenience, the state space of  $X$  can be taken to be, without loss of generality, the set of unit vectors  $\{e_1, e_2, \dots, e_N\}$ , where  $e_i = (0, \dots, 0, 1, 0, \dots, 0)' \in \mathbb{R}^N$  with the “1” in the  $i$ th entry and where the superscript  $'$  represents the transpose of a row vector. This is called the *canonical representation* of the state space of the Markov chain  $X$ . Suppose the process  $X$  is homogeneous in time. Let  $A$  denote the constant rate matrix of the Markov chain  $X$ . Then, using the canonical representation of the state space, the dynamics of the Markov chain  $X$  have the following *semi-martingale representation* (see Elliott et al. (1995))

$$dX_t = AX_t dt + dM_t, \tag{4.1}$$

where  $(M_t)_{t \geq 0}$  is an  $\mathbb{R}^N$ -valued martingale with respect to the natural filtration generated by  $X$ .

We suppose the short (instantaneous) interest rate  $r$  depends on the state  $X$  of the economy, so that

$$r_t = r(X_t) = \langle r, X_t \rangle,$$

where  $r = (r_1, r_2, \dots, r_N)' \in \mathbb{R}^N$ , and each  $r_i$  is a constant,  $1 \leq i \leq N$ .

Suppose  $W$  is a standard Brownian motion on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Consider a firm whose asset value  $V$  follows a **regime switching** geometric Brownian motion<sup>1</sup>

$$\begin{aligned} \frac{dV_t}{V_t} &= (r_t - a_t)dt + \sigma_t dW_t \\ &= \langle r - a, X_t \rangle dt + \langle \sigma, X_t \rangle dW_t, \end{aligned}$$

where  $a_t = a(X_t)$  is the payout rate at time  $t$ , which is also switching with the Markov chain  $X$ . Thus, there exists a vector  $a := (a_1, \dots, a_N)$  such that  $a_t = \langle a, X_t \rangle$ .

Let  $\mu_t := r_t - a_t = \langle r - a, X_t \rangle$ . Then

$$\frac{dV_t}{V_t} = \mu_t dt + \sigma_t dW_t.$$

Here, the drift  $\mu_t = \mu(X_t)$  and volatility  $\sigma_t = \sigma(X_t)$  also depend on the state  $X_t$  of the economy. That is, there are vectors  $\mu = (\mu_1, \dots, \mu_N)'$  and  $\sigma = (\sigma_1, \dots, \sigma_N)'$  such that  $\mu_t = \langle \mu, X_t \rangle$  (i.e.,  $\mu = r - a$ ) and  $\sigma_t = \langle \sigma, X_t \rangle$ .

The stochastic process  $V$  is assumed to be unaffected by the financial structure of the firm. This is consistent with bond covenants that restrict firms from selling assets. Brennan and Schwartz (1978), Leland (1994), Leland and Toft (1996) etc., also make this assumption, although Merton (1974) does not.

---

<sup>1</sup>As in previous contributions, e.g. Leland (1994) and Leland and Toft (1996), we assume, without loss of generality, the dynamics are in the risk-neutral measure.

#### 4.2.2 Coupled PDE for General Cash Flow Stream

We first derive the coupled PDE for the value of the coupon bond. We can obtain the following theorem for a general cash flow stream. Let  $S$  be a process, which models the dynamics of an underlying risky asset, with

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t,$$

where  $\mu_t = \mu(X_t) = \langle \mu, X \rangle$ ,  $\sigma_t = \sigma(X_t) = \langle \sigma, X \rangle$ .

Consider a general contingent claim which has a continuous dividend stream  $g(t, S_t, X_t)$  at time  $t$ , and a terminal payment  $f(S_T, X_T)$  at maturity  $T$ . The value of the contingent claim at time  $t$  is

$$P_i(t, S) = \mathbb{E} \left[ \int_t^T e^{-\int_t^u r_s ds} g(u, S_u, X_u) du + e^{-\int_t^T r_u du} f(S_T, X_T) | S_t = S, X_t = e_i \right].$$

We can write this in an equivalent form

$$P(t, S, X) = \mathbb{E} \left[ \int_t^T e^{-\int_t^u r_v dv} g(u, S_u, X_u) du + e^{-\int_t^T r_u du} f(S_T, X_T) | S_t = S, X_t = X \right],$$

Here  $X = X_t$  takes values in  $\{e_1, \dots, e_N\}$ .

**Theorem 4.2.1** *Under some technical conditions, the function  $P(t, S, X) := P(t, S_t, X_t)$  is the solution of the coupled PDE*

$$P_t + \frac{1}{2} \sigma^2 S^2 P_{SS} + \mu S P_S - rV + g + \langle P, AX \rangle = 0,$$

$$P(T, x, X) = f(x, X_T).$$

*Proof.* Let  $P(t, S, X) := P(t, S_t, X_t)$ . Suppose  $P(t, S, X)$  satisfies the coupled PDE. We consider  $e^{-\int_t^s r_u du} P(s, S_s, X_s)$ . Using Itô's rule, we have

$$\begin{aligned} d \left( e^{-\int_t^s r_u du} P(s, S_s, X_s) \right) &= e^{-\int_t^s r_u du} \left( P_t + \frac{1}{2} \sigma^2 S^2 P_{SS} + \mu S P_S - rV + \langle P, AX \rangle \right) dt \\ &\quad + [\dots] dW_t + e^{-\int_0^t r_u du} \langle P, dM_t \rangle. \end{aligned}$$

Integrating from  $t$  to  $T$  (letting variable  $s$  range from  $t$  to  $T$ ), we have

$$\begin{aligned}
e^{-\int_t^T r_u du} P(T, S_T, X_T) - e^{-\int_t^t r_u du} P(t, S_t, X_t) &= \int_t^T e^{-\int_t^s r_u du} \left( P_t + \frac{1}{2} \sigma^2 S^2 P_{SS} + \mu S P_S - rV + \langle P, AX \rangle \right) dt \\
&\quad + [\dots] dW_t + e^{-\int_0^t r_u du} \langle P, dM_t \rangle \\
&= \int_t^T e^{-\int_t^v r_u du} (-g(v, S_v, X_v)) dv + \int_t^T [\dots] dW_t + \int_t^T e^{-\int_0^t r_u du} \langle P, dM_t \rangle.
\end{aligned}$$

(from coupled PDE)

Using the boundary condition in the coupled PDE, we see that

$$e^{-\int_t^T r_u du} f(S_T, X_T) + \int_t^T e^{-\int_t^s r_u du} g(v, S_v, X_v) dv - [\dots] dW_t = P(t, S_t, X_t)$$

Under technical conditions, the conditional expectation  $\mathbb{E}[\dots dW_t | \mathcal{F}_t]$  of the  $dW_t$ -term vanishes. Thus, taking conditional expectation  $\mathbb{E}[\cdot | \mathcal{F}_t]$ , we see that,

$$\begin{aligned}
P(t, S, X) &= \mathbb{E} \left[ e^{-\int_t^T r_u du} f(S_T, X_T) + \int_t^T e^{-\int_t^s r_u du} g(v, S_v, X_v) dv | \mathcal{F}_t \right] \\
&= \mathbb{E} \left[ e^{-\int_t^T r_u du} f(S_T, X_T) + \int_t^T e^{-\int_t^s r_u du} g(v, S_v, X_v) dv | S_t = S, X_t = X \right].
\end{aligned}$$

This is the definition of  $P(t, x)$ . □

### 4.3 Simple Optimal Capital Structure with Regime Switching

Consider a firm's debt. Suppose the debt of the firm is in the form of a claim on the firm that continuously pays a nonnegative coupon at rate  $C$  at every instant of time. We assume that the firm finances the net cost of the coupon by issuing additional equity, as postulated in Brennan and Schwartz (1978), Leland (1994), Leland and Toft (1996) etc.

The total value of the firm's assets is, therefore,  $V_t$ , and the value of the debt (bond) is denoted  $D_t$ . Therefore the value of the firm's equity is  $E_t = V_t - D_t$ . Note that the value of debt,  $D_t$ , actually depends on  $V_t$ , and therefore, it is more precise to write  $D_t = D(t, V_t)$ .

To derive closed-form explicit solutions, we examine debts that depend on underlying firm value  $V$  but are otherwise time independent. More precisely, we consider perpetual bonds (*consol bonds*), namely bonds promising a fixed coupon with no final maturity date.

Consol debt promises a perpetual coupon payment  $C$ , whose level remains constant unless the firm declares bankruptcy. As the bond is perpetual, it is time-independent, and thus we can write  $D_t = D(V_t)$ . Suppose  $v_i$  is the critical level of the value of the firm,  $V$ , at which bankruptcy is declared, when the initial state of  $X$  is  $X_0 = e_i$ ,  $i = 1, \dots, N$ . The bankruptcy time determined in such a way is  $\tau^* = \inf\{t \geq 0 : V_t \leq \bar{v}\} = \inf\{t \geq 0 : V_t < \bar{v}\}$ . To find the actual value of  $v_i$ , and obtain closed-form explicit solutions for optimal capital structure, we consider the coupled PDE.

From Theorem 4.2.1, we see that the coupled ODEs for the pricing function  $P(x)$  of a consol bond are

$$\frac{1}{2}x^2\sigma^2P''(x) + \mu xP'(x) - rP(x) + c + \langle P, AX \rangle = 0, \quad (4.2)$$

subject to the boundary condition

$$P(v_i) = \min(v_i, c/r)$$

$$\lim_{x \rightarrow \infty} P(x) = c/r.$$

For simplicity, we examine the case where  $X$  has two states, i.e., the economy is shifting between two regimes. Then, the rate matrix  $A$  for the Markov chain  $X$  is of form<sup>2</sup>

$$A = \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{bmatrix}' = \begin{bmatrix} -\lambda_1 & \lambda_2 \\ \lambda_1 & -\lambda_2 \end{bmatrix}.$$

We are able to obtain a *closed-form* analytical expression for the firm's optimal capital structure.

**Theorem 4.3.1** *Let  $v_1$  and  $v_2$  be the optimal default barrier, and  $P_1(x)$  and  $P_2(x)$  the optimal (minimum) value of the firm's defaultable debt, when  $X_t = e_1$  or  $e_2$  respectively. Then*

- 1) *the optimal default barriers  $v_1$  and  $v_2$  are given by (see equations (4.23), (4.24),*

---

<sup>2</sup>Note that in our setting  $A$  is the transpose of the usual generator matrix of a Markov chain.

(4.25)

$$\begin{bmatrix} v_1^{-\gamma_1} & 0 \\ 0 & v_1^{-\gamma_2} \end{bmatrix} F_1(v_1) = \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} F_2(x_2),$$

where

$$F_1(v_1) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} v_1 - \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} v_1 - \frac{c}{r_1 + \lambda_1} \\ v_1 - v_1 \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} \end{bmatrix},$$

and

$$F_2(v_2) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \left( \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix} - \begin{bmatrix} \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} v_2 + \frac{c}{r_1 + \lambda_1} - P_1^* \\ v_2 \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} \end{bmatrix} \right)$$

2) The optimal (minimum) value of the firm's defaultable debt is

$$P_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} + \frac{c(\lambda_1 - \lambda_2)}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} x + \frac{c}{r_1 + \lambda_1} & \text{if } v_1 \leq x \leq v_2 \\ x & \text{if } x \leq v_1, \end{cases}$$

and

$$P_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} + \frac{c((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ x & \text{if } x \leq v_2, \end{cases}$$

where  $\beta_1, \beta_2$  are the two negative roots of equation (4.9), and  $\gamma_1, \gamma_2$  are the two roots of equation (4.14).

3) The coefficients  $A_i, B_i,$  and  $C_i$  are given in equations (4.26), (4.27), and (4.28).

*Proof.* See Appendix. □

#### 4.4 Bankruptcy Costs and Tax Effect

In this section we follow Leland (1994) and consider the tax benefits and bankruptcy costs.

#### 4.4.1 Bankruptcy Costs

We assume that the firm defaults as soon as the firm's value  $V_t$  hits the default barrier  $\bar{v}$ . We extend the previous section to allow for bankruptcy costs. Suppose that at the time of default the bondholders receive a recovery payment of  $\alpha\bar{v}$ , where  $\alpha \in [0, 1]$ . The fraction  $(1 - \alpha)\bar{v}$  represents bankruptcy or reorganization costs.

To facilitate our smooth pasting technique, we regard the bankruptcy costs as a claim or security which pays no coupon, but has payment equal to the bankruptcy costs  $(1 - \alpha)\bar{v}$  at the time of default. This security has current value, denoted  $B(V)$ , that reflects the market value of a claim to  $(1 - \alpha)\bar{v}$  should bankruptcy occur.

#### 4.4.2 Tax Benefits

Now consider the value of tax benefits associated with debt financing. These benefits resemble a security that pays a constant coupon equal to the tax-sheltering value of interest payments  $C' = \theta C$  as long as the firm is solvent and pays nothing at the time of bankruptcy. This security's value, denoted  $T(V)$ , equals the value of the tax benefit of debt.

#### 4.4.3 Value of the Firm

The total value of the firm, denoted  $v(V)$ , reflects three terms: the firm's asset value, plus the value of the tax deduction of coupon payments, less the value of bankruptcy costs. Thus

$$v(V) = V + T(V) - B(V).$$

As in the previous section, the firm issues debt, which is of the form of a perpetual (consol) bond. Then the value of equity is the total value of the firm less the value of debt

$$\begin{aligned} E(V) &= v(V) - D(V) = V + T(V) - B(V) - D(V) \\ &= V - (D(V) + B(V) - T(V)). \end{aligned}$$

The equity holders' objective is to maximize the value of the equity, or equivalently, minimize the value of  $D(V) + B(V) - T(V)$ .

#### 4.4.4 A New Security for Smooth Pasting

To make use of the smooth pasting technique, as in the previous section, we consider a new security (claim), defined to be

$$D'(V) = D(V) + B(V) - T(V).$$

Thus, the security  $D'(V)$  represents a continuous coupon payment  $C$  at any time instant until the time of default, plus a payment equal to the bankruptcy costs  $(1 - \alpha)\bar{v}$  at the time of default, less a continuous constant coupon equal to the tax-sheltering value of coupon payments  $c' = \theta c$  as long as the firm is solvent. In other words, the security  $D'(V)$  has a continuous coupon payment  $c - c' = (1 - \theta)c$  at any time instant as long as the firm is solvent, plus a payment equal to the bankruptcy costs  $(1 - \alpha)\bar{v}$  at the time of default.

#### 4.4.5 Optimal Capital Structure with Tax and Bankruptcy Costs under Regime Switching

Let  $P(x)$  be the value of the claim  $D'$  when  $V_0 = x$ . Using Theorem 4.2.1 in Section 2, we have the following coupled ODEs for the pricing function  $P(x)$  of a consol bond:

$$\frac{1}{2}x^2\sigma^2P''(x) + \mu xP'(x) - rP(x) + c - c' + \langle P, AX \rangle = 0, \quad (4.3)$$

subject to the boundary condition

$$P(v_i) = \min((1 - \alpha)v_i, (c - c')/r)$$

$$\lim_{x \rightarrow \infty} P(x) = (c - c')/r.$$

For simplicity, we examine the case where  $X$  has two states, i.e., the economy is shifting between two regimes. We are able to obtain a *closed-form* analytical solution for the firm's optimal capital structure.

**Theorem 4.4.1** *Let  $v_1$  and  $v_2$  be the optimal default barrier, and  $P_1(x)$  and  $P_2(x)$  the optimal (minimum) value of the firm's defaultable debt, when  $X_t = e_1$  or  $e_2$  respectively. Then*



1) the optimal default barriers  $v_1$  and  $v_2$  are given by (see equations (4.48), (4.49), (4.50))

$$\begin{bmatrix} v_1^{-\gamma_1} & 0 \\ 0 & v_1^{-\gamma_2} \end{bmatrix} F_1(v_1) = \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} F_2(x_2),$$

where

$$F_1(v_1) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_1 - \frac{\lambda_1(1-\alpha)}{r_1+\lambda_1-\mu_1}v_1 - \frac{c-c'}{r_1+\lambda_1} \\ (1-\alpha)v_1 - v_1 \frac{\lambda_1(1-\alpha)}{r_1+\lambda_1-\mu_1} \end{bmatrix},$$

and

$$F_2(v_2) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \left( \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_2 - Q_2^* \\ (1-\alpha)v_2 \end{bmatrix} - \begin{bmatrix} \frac{\lambda_1(1-\alpha)}{r_1+\lambda_1-\mu_1}v_2 + \frac{c-c'}{r_1+\lambda_1} - Q_1^* \\ (1-\alpha)v_2 \frac{\lambda_1}{r_1+\lambda_1-\mu_1} \end{bmatrix} \right),$$

and  $Q_1^*$  and  $Q_2^*$  are given in (4.35).

2) The optimal (minimum) value of the firm's defaultable debt is

$$P_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} + \frac{c(\lambda_1-\lambda_2)}{r_2\lambda_1-r_1\lambda_2} & \text{if } x > v_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \frac{\lambda_1}{r_1+\lambda_1-\mu_1}x + \frac{c}{r_1+\lambda_1} & \text{if } v_1 \leq x \leq v_2 \\ x & \text{if } x \leq v_1, \end{cases}$$

and

$$P_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} + \frac{c((r_1+\lambda_1)-(r_2+\lambda_2))}{r_2\lambda_1-r_1\lambda_2} & \text{if } x > v_2, \\ x & \text{if } x \leq v_2, \end{cases}$$

where  $\beta_1, \beta_2$  are the two negative roots of equation (4.34), and  $\gamma_1, \gamma_2$  are the two roots of equation (4.39).

3) The coefficients  $A_i, B_i,$  and  $C_i$  are given in equations (4.51), (4.52), and (4.53).

*Proof.* See Appendix. □

## 4.5 Concluding Remarks

We investigated the optimal capital structure of a corporate when the dynamics of the assets (both growth rate and volatility) change following different states of the economy.

We consider two models. The first model considers the case when the firm is not facing tax benefit and bankruptcy costs, with regime switching dynamics. This extends the Black and Cox (1976) model to allow for regime switching risk.

The second model incorporates tax benefit and bankruptcy costs with regime switching dynamics. This is more realistic, and is an extension of the Leland (1994) model with regime switching risk.

We obtain **closed-form** analytic solutions for the optimal capital structures and default barriers for both models.

## Appendix

*Proof of Theorem 4.3.1.* Note that

$$A = \begin{bmatrix} -\lambda_1 & \lambda_2 \\ \lambda_1 & -\lambda_2 \end{bmatrix}.$$

Then the coupled homogeneous ODEs reduce to the following system

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(P_2(x) - P_1(x)) = 0, \\ P_1(v_1) = \min(v_1, c/r), \\ \lim_{x \rightarrow \infty} P_1(x) = c/r. \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_2^2P_2''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c + \lambda_2(P_1(x) - P_2(x)) = 0, \\ P_2(v_2) = \min(v_2, c/r), \\ \lim_{x \rightarrow \infty} P_2(x) = c/r. \end{array} \right.$$

For simplicity, we further assume that  $v_1 \leq v_2 \leq c/r$ , as if  $v_i \geq c/r$ , then  $P_i(v_i) = c/r$ , and moreover  $P_i(x) = c/r$  for all  $x \geq v_i$  (as the equity holders will declare bankruptcy immediately at time 0). Note that the value of the default-free perpetual bond is  $c/r$ . Intuitively, if the barrier  $v_i \geq c/r$ , the equity holders choose to default immediately at time 0. To obtain the optimal values  $v_i$ , we also impose the smooth pasting conditions at  $v_i$ . Therefore

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(P_2(x) - P_1(x)) = 0 \\ P_1(v_1) = v_1 \\ \lim_{x \rightarrow \infty} P_1(x) = c/r, \\ P_1'(x)|_{x=v_1} = 1. \end{array} \right. \quad (4.4)$$

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_2^2P_2''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c + \lambda_2(P_1(x) - P_2(x)) = 0 \\ P_2(v_2) = v_2 \\ \lim_{x \rightarrow \infty} P_2(x) = c/r, \\ P_2'(x)|_{x=v_2} = 1. \end{array} \right. \quad (4.5)$$

Therefore, when  $v_1 \leq v_2 \leq c/r$ , we can write the coupled ODEs as follows. For  $x \in [0, v_1]$ , we have

$$P_1(x) = P_2(x) = x. \quad (4.6)$$

For  $x \in [v_1, v_2]$ , we have

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(P_2(x) - P_1(x)) = 0 \\ P_2(x) = x. \end{array} \right. \quad (4.7)$$

For  $x \in [v_2, \infty)$ , we have

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(P_2(x) - P_1(x)) = 0, \\ \frac{1}{2}x^2\sigma_2^2P_2''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c + \lambda_2(P_1(x) - P_2(x)) = 0. \end{array} \right. \quad (4.8)$$

Now (5.16) has an characteristic function

$$g_1(\beta)g_2(\beta) = \lambda_1\lambda_2, \quad (4.9)$$

where

$$g_1(\beta) = \lambda_1 + r_1 - \mu_1\beta - \frac{1}{2}\sigma_1^2\beta(\beta - 1),$$

$$g_2(\beta) = \lambda_2 + r_2 - \mu_2\beta - \frac{1}{2}\sigma_2^2\beta(\beta - 1).$$

This characteristic function has four distinct roots  $\beta_1 < \beta_2 < 0 < \beta_3 < \beta_4$ . To obtain a particular solution of (5.16), we consider

$$\begin{cases} -r_1P_1^* + c + \lambda_1(P_2^* - P_1^*) = 0 \\ -r_2P_2^* + c + \lambda_2(P_1^* - P_2^*) = 0. \end{cases}$$

This system reduces to

$$\begin{bmatrix} -(r_1 + \lambda_1) & \lambda_1 \\ -(r_2 + \lambda_2) & \lambda_2 \end{bmatrix} \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix} = \begin{bmatrix} -c \\ -c \end{bmatrix},$$

and hence

$$\begin{aligned} \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix} &= \begin{bmatrix} -(r_1 + \lambda_1) & \lambda_1 \\ -(r_2 + \lambda_2) & \lambda_2 \end{bmatrix}^{-1} \begin{bmatrix} -c \\ -c \end{bmatrix} \\ &= \frac{1}{r_2\lambda_1 - r_1\lambda_2} \begin{bmatrix} \lambda_2 & -\lambda_1 \\ r_2 + \lambda_2 & -(r_1 + \lambda_1) \end{bmatrix} \begin{bmatrix} -c \\ -c \end{bmatrix} \\ &= \frac{1}{r_2\lambda_1 - r_1\lambda_2} \begin{bmatrix} c(\lambda_1 - \lambda_2) \\ c((r_1 + \lambda_1) - (r_2 + \lambda_2)) \end{bmatrix}, \end{aligned}$$

Thus a particular solution of (5.16) is

$$P_1^* = \frac{c(\lambda_1 - \lambda_2)}{r_2\lambda_1 - r_1\lambda_2}$$

$$P_2^* = \frac{c((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2\lambda_1 - r_1\lambda_2}. \quad (4.10)$$

Therefore, the general form of the solution to (5.16) is

$$\begin{aligned} P_1(x) &= P_1^* + A_1x^{\beta_1} + A_2x^{\beta_2} + A_3x^{\beta_3} + A_4x^{\beta_4}, \\ P_2(x) &= P_2^* + B_1x^{\beta_1} + B_2x^{\beta_2} + B_3x^{\beta_3} + B_4x^{\beta_4}, \end{aligned}$$

with  $B_i = l_i A_i$  and  $l_i = l(\beta_i) = g_1(\beta_i)/\lambda_1 = \lambda_2/g_2(\beta_i)$ .

When  $x \rightarrow \infty$ ,  $P_1(x)$  and  $P_2(x)$  are both bounded. Thus  $A_3 = A_4 = B_3 = B_4 = 0$ , and the solution is

$$\begin{aligned} P_1(x) &= P_1^* + A_1x^{\beta_1} + A_2x^{\beta_2}, \\ P_2(x) &= P_2^* + B_1x^{\beta_1} + B_2x^{\beta_2}. \end{aligned} \tag{4.11}$$

Next we solve (5.15). The first equation is

$$\frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(x - P_1(x)) = 0. \tag{4.12}$$

This is an inhomogeneous equation, and thus the solution can be written as

$$P_1(x) = C_1x^{\gamma_1} + C_2x^{\gamma_2} + \phi(x), \tag{4.13}$$

where  $\phi(x)$  is a particular solution and  $\gamma_1$  and  $\gamma_2$  are the two roots of

$$\frac{1}{2}\sigma_1^2\gamma(\gamma - 1) + \mu_1\gamma - r_1 - \lambda_1 = 0. \tag{4.14}$$

To obtain  $\phi(x)$ , we assume that  $\phi(x) = ax + b$  and substitute it to equation (5.22). This yields

$$a\mu_1x - r_1(ax + b) + c - \lambda_1(ax + b) + \lambda_1x = 0,$$

or

$$x((\mu_1 - r_1 - \lambda_1)a + \lambda_1) + c - r_1b - \lambda_1b = 0.$$

Thus  $a = \lambda_1/(r_1 + \lambda_1 - \mu_1)$ ,  $b = c/(r_1 + \lambda_1)$ , and thus

$$\phi(x) = \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1}x + \frac{c}{r_1 + \lambda_1}. \tag{4.15}$$

Now we solve for the coefficients  $A_i$ ,  $B_i$ ,  $C_i$  and the optimal values  $v_1$  and  $v_2$ . Using the boundary and smooth pasting conditions for  $P_2(x)$  (see(5.13), (5.16)) at  $v_2$  with  $x \in [v_2, \infty)$ ,

$$\begin{cases} P_2^* + B_1 v_2^{\beta_1} + B_2 v_2^{\beta_2} = v_2, \\ \beta_1 B_1 v_2^{\beta_1 - 1} + \beta_2 B_2 v_2^{\beta_2 - 1} = 1. \end{cases}$$

or

$$\begin{cases} l_1 A_1 v_2^{\beta_1} + l_2 A_2 v_2^{\beta_2} = v_2 - P_2^*, \\ \beta_1 l_1 A_1 v_2^{\beta_1} + \beta_2 l_2 A_2 v_2^{\beta_2} = v_2. \end{cases} \quad (4.16)$$

Similarly, using the boundary and smooth pasting conditions for  $P_1(x)$  (see (5.12), (5.15)) at  $v_2$  with  $x \in [v_1, v_2]$ ,

$$\begin{cases} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} = C_1 v_2^{\gamma_1} + C_2 v_2^{\gamma_2} + \phi(v_2), \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} = \gamma_1 C_1 v_2^{\gamma_1} + \gamma_2 C_2 v_2^{\gamma_2} + v_2 \phi'(v_2). \end{cases} \quad (4.17)$$

Using the boundary and smooth pasting conditions for  $P_1(x)$  (see (5.12), (5.15)) at  $v_1$  with  $x \in [v_1, v_2]$

$$\begin{cases} C_1 v_1^{\gamma_1} + C_2 v_1^{\gamma_2} + \phi(v_1) = v_1, \\ \gamma_1 C_1 v_1^{\gamma_1} + \gamma_2 C_2 v_1^{\gamma_2} + v_1 \phi'(v_1) = v_1. \end{cases} \quad (4.18)$$

Combining these equations, we can obtain the solutions  $v_1$  and  $v_2$ . From (5.28), we have

$$F_1(v_1) := \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} v_1 - \phi(v_1) \\ v_1 - v_1 \phi'(v_1) \end{bmatrix} = \begin{bmatrix} C_1 v_1^{\gamma_1} \\ C_2 v_1^{\gamma_2} \end{bmatrix}. \quad (4.19)$$

From (5.26), we have

$$\begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix} = \begin{bmatrix} A_1 v_2^{\beta_1} \\ A_2 v_2^{\beta_2} \end{bmatrix}. \quad (4.20)$$

Then, using (5.27) and (4.20), we have

$$\begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} A_1 v_2^{\beta_1} \\ A_2 v_2^{\beta_2} \end{bmatrix} = \begin{bmatrix} C_1 v_2^{\gamma_1} + C_2 v_2^{\gamma_2} + \phi(v_2) - P_2^* \\ \gamma_1 C_1 v_2^{\gamma_1} + \gamma_2 C_2 v_2^{\gamma_2} + v_2 \phi'(v_2) \end{bmatrix}. \quad (4.21)$$

This equation can be rewritten as

$$\begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix} - \begin{bmatrix} \phi(v_2) - P_1^* \\ v_2 \phi'(v_2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix} \begin{bmatrix} C_1 v_2^{\gamma_1} \\ C_2 v_2^{\gamma_2} \end{bmatrix}.$$

Therefore

$$F_2(v_2) := \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \left( \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix} - \begin{bmatrix} \phi(v_2) - P_1^* \\ v_2 \phi'(v_2) \end{bmatrix} \right) = \begin{bmatrix} C_1 v_2^{\gamma_1} \\ C_2 v_2^{\gamma_2} \end{bmatrix}. \quad (4.22)$$

Hence, from (4.19) and (4.22) we have the equation

$$\begin{bmatrix} v_1^{-\gamma_1} & 0 \\ 0 & v_1^{-\gamma_2} \end{bmatrix} F_1(v_1) = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} F_2(v_2). \quad (4.23)$$

In particular, for  $\phi(x)$  of the form in (5.25), we have from (4.19) and (4.22)

$$F_1(v_1) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} v_1 - \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} v_1 - \frac{c}{r_1 + \lambda_1} \\ v_1 - v_1 \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} \end{bmatrix}, \quad (4.24)$$

and

$$F_2(v_2) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \left( \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix} - \begin{bmatrix} \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} v_2 + \frac{c}{r_1 + \lambda_1} - P_1^* \\ v_2 \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} \end{bmatrix} \right). \quad (4.25)$$

Substituting  $F_1(v_1)$  and  $F_2(v_2)$  into (4.23), we can obtain the values for  $v_1$  and  $v_2$ . Now we derive coefficients  $A_i$ ,  $B_i$ , and  $C_i$ . From (5.26), the coefficients  $A_1$  and  $A_2$  are give by

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} l_1 v_2^{\beta_1} & l_2 v_2^{\beta_2} \\ \beta_1 l_1 v_2^{\beta_1} & \beta_2 l_2 v_2^{\beta_2} \end{bmatrix}^{-1} \begin{bmatrix} v_2 - P_2^* \\ v_2 \end{bmatrix}, \quad (4.26)$$

and

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} l_1 A_1 \\ l_2 A_2 \end{bmatrix}. \quad (4.27)$$

From (5.28), we have

$$\begin{aligned} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} &= \begin{bmatrix} v_1^{\gamma_1} & v_1^{\gamma_2} \\ \gamma_1 v_1^{\gamma_1} & \gamma_2 v_1^{\gamma_2} \end{bmatrix}^{-1} \begin{bmatrix} v_1 - \phi(v_1) \\ v_1 - v_1 \phi'(v_1) \end{bmatrix} \\ &= \begin{bmatrix} v_1^{\gamma_1} & v_1^{\gamma_2} \\ \gamma_1 v_1^{\gamma_1} & \gamma_2 v_1^{\gamma_2} \end{bmatrix}^{-1} \begin{bmatrix} v_1 - \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} v_1 - \frac{c}{r_1 + \lambda_1} \\ v_1 - v_1 \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} \end{bmatrix}. \end{aligned} \quad (4.28)$$

With these coefficients, the value functions  $P_1(x)$  and  $P_2(x)$  become

$$P_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} + \frac{c(\lambda_1 - \lambda_2)}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \phi(x) & \text{if } v_1 \leq x \leq v_2 \\ x & \text{if } x \leq v_1, \end{cases}$$

and

$$P_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} + \frac{c((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ x & \text{if } x \leq v_2, \end{cases}$$

where

$$\phi(x) = \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} x + \frac{c}{r_1 + \lambda_1}.$$

□

*Proof of Theorem 4.4.1.* The coupled homogeneous ODEs reduce to the following system

$$\begin{cases} \frac{1}{2} x^2 \sigma_1^2 P_1''(x) + \mu_1 x P_1'(x) - r_1 P_1(x) + c - c' + \lambda_1 (P_2(x) - P_1(x)) = 0, \\ P_1(v_1) = \min((1 - \alpha)v_i, (c - c')/r), \\ \lim_{x \rightarrow \infty} P_1(x) = (c - c')/r. \end{cases}$$



$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_2^2P_2''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c - c' + \lambda_2(P_1(x) - P_2(x)) = 0, \\ P_2(v_2) = \min((1 - \alpha)v_i, (c - c')/r), \\ \lim_{x \rightarrow \infty} P_2(x) = (c - c')/r. \end{array} \right.$$

For simplicity, we assume that  $(1 - \alpha)v_1 \leq (1 - \alpha)v_2 \leq (c - c')/r$ , as if  $(1 - \alpha)v_i \geq (c - c')/r$ , then  $P_i(v_i) = \min((1 - \alpha)v_i, (c - c')/r) = (c - c')/r$ , and moreover  $P_i(x) = (c - c')/r$  for all  $x \geq v_i$ . Note that the value of the default-free perpetual bond is  $(c - c')/r$ . Intuitively, if the barrier  $v_i \geq (c - c')/r$ , the equity holders choose to default immediately. To obtain the optimal values  $v_i$ , we also impose the smooth pasting conditions at  $v_i$ . Therefore

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c - c' + \lambda_1(P_2(x) - P_1(x)) = 0 \\ P_1(v_1) = (1 - \alpha)v_1 \\ \lim_{x \rightarrow \infty} P_1(x) = (c - c')/r, \\ P_1'(x)|_{x=v_1} = 1 - \alpha. \end{array} \right. \quad (4.29)$$

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_2^2P_2''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c - c' + \lambda_2(P_1(x) - P_2(x)) = 0 \\ P_2(v_2) = (1 - \alpha)v_2 \\ \lim_{x \rightarrow \infty} P_2(x) = (c - c')/r, \\ P_2'(x)|_{x=v_2} = 1 - \alpha. \end{array} \right. \quad (4.30)$$

Therefore, when  $v_1 \leq v_2 \leq (c - c')/r$ , we can write the coupled ODEs as follows. For  $x \in [0, v_1]$ , we have

$$P_1(x) = P_2(x) = (1 - \alpha)x. \quad (4.31)$$

For  $x \in [v_1, v_2]$ , we have

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c - c' + \lambda_1(P_2(x) - P_1(x)) = 0 \\ P_2(x) = (1 - \alpha)x. \end{array} \right. \quad (4.32)$$

For  $x \in [v_2, \infty)$ , we have

$$\begin{cases} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c - c' + \lambda_1(P_2(x) - P_1(x)) = 0, \\ \frac{1}{2}x^2\sigma_2^2P_1''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c - c' + \lambda_2(P_1(x) - P_2(x)) = 0. \end{cases} \quad (4.33)$$

Now (4.33) has an characteristic function

$$g_1(\beta)g_2(\beta) = \lambda_1\lambda_2, \quad (4.34)$$

where

$$\begin{aligned} g_1(\beta) &= \lambda_1 + r_1 - \mu_1\beta - \frac{1}{2}\sigma_1^2\beta(\beta - 1), \\ g_2(\beta) &= \lambda_2 + r_2 - \mu_2\beta - \frac{1}{2}\sigma_2^2\beta(\beta - 1). \end{aligned}$$

This characteristic function has four distinct roots  $\beta_1 < \beta_2 < 0 < \beta_3 < \beta_4$ . To obtain a particular solution of (5.16), we consider

$$\begin{cases} -r_1Q_1^* + c - c' + \lambda_1(Q_2^* - Q_1^*) = 0 \\ -r_2Q_2^* + c - c' + \lambda_2(Q_1^* - Q_2^*) = 0. \end{cases}$$

This system reduces to

$$\begin{bmatrix} -(r_1 + \lambda_1) & \lambda_1 \\ -(r_2 + \lambda_2) & \lambda_2 \end{bmatrix} \begin{bmatrix} Q_1^* \\ Q_2^* \end{bmatrix} = \begin{bmatrix} c' - c \\ c' - c \end{bmatrix},$$

and hence

$$\begin{aligned} \begin{bmatrix} Q_1^* \\ Q_2^* \end{bmatrix} &= \begin{bmatrix} -(r_1 + \lambda_1) & \lambda_1 \\ -(r_2 + \lambda_2) & \lambda_2 \end{bmatrix}^{-1} \begin{bmatrix} c' - c \\ c' - c \end{bmatrix} \\ &= \frac{1}{r_2\lambda_1 - r_1\lambda_2} \begin{bmatrix} \lambda_2 & -\lambda_1 \\ r_2 + \lambda_2 & -(r_1 + \lambda_1) \end{bmatrix} \begin{bmatrix} c' - c \\ c' - c \end{bmatrix} \\ &= \frac{1}{r_2\lambda_1 - r_1\lambda_2} \begin{bmatrix} (c - c')(\lambda_1 - \lambda_2) \\ (c - c')((r_1 + \lambda_1) - (r_2 + \lambda_2)) \end{bmatrix}, \end{aligned}$$

Thus a particular solution of (5.16) is

$$\begin{aligned} Q_1^* &= \frac{(c - c')(\lambda_1 - \lambda_2)}{r_2\lambda_1 - r_1\lambda_2} \\ Q_2^* &= \frac{(c - c')((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2\lambda_1 - r_1\lambda_2}. \end{aligned} \quad (4.35)$$

Therefore the general form of the solution to (4.33) is

$$\begin{aligned} P_1(x) &= Q_1^* + A_1x^{\beta_1} + A_2x^{\beta_2} + A_3x^{\beta_3} + A_4x^{\beta_4}, \\ P_2(x) &= Q_2^* + B_1x^{\beta_1} + B_2x^{\beta_2} + B_3x^{\beta_3} + B_4x^{\beta_4}, \end{aligned}$$

with  $B_i = l_i A_i$  and  $l_i = l(\beta_i) = g_1(\beta_i)/\lambda_1 = \lambda_2/g_2(\beta_i)$ .

When  $x \rightarrow \infty$ ,  $P_1(x)$  and  $P_2(x)$  are both bounded. Thus  $A_3 = A_4 = B_3 = B_4 = 0$ , and the solution is

$$\begin{aligned} P_1(x) &= Q_1^* + A_1x^{\beta_1} + A_2x^{\beta_2}, \\ P_2(x) &= Q_2^* + B_1x^{\beta_1} + B_2x^{\beta_2}. \end{aligned} \quad (4.36)$$

Next we solve (4.32). The first equation is

$$\frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c - c' + \lambda_1((1 - \alpha)x - P_1(x)) = 0. \quad (4.37)$$

This is an inhomogeneous equation, and thus the solution can be written as

$$P_1(x) = C_1x^{\gamma_1} + C_2x^{\gamma_2} + \phi(x), \quad (4.38)$$

where  $\phi(x)$  is a particular solution and  $\gamma_1$  and  $\gamma_2$  are the two roots of

$$\frac{1}{2}\sigma_1^2\gamma(\gamma - 1) + \mu_1\gamma - r_1 - \lambda_1 = 0. \quad (4.39)$$

To obtain  $\phi(x)$ , we assume that  $\phi(x) = ax + b$  and substitute it to equation (4.37). This yields

$$a\mu_1x - r_1(ax + b) + c - c' + \lambda_1(1 - \alpha)x - \lambda_1(ax + b) = 0,$$

or

$$x((\mu_1 - r_1 - \lambda_1)a + \lambda_1(1 - \alpha)) + c - c' - r_1b - \lambda_1b = 0.$$

Thus  $a = \lambda_1(1 - \alpha)/(r_1 + \lambda_1 - \mu_1)$ ,  $b = (c - c')/(r_1 + \lambda_1)$ , and thus

$$\phi(x) = \frac{\lambda_1(1 - \alpha)}{r_1 + \lambda_1 - \mu_1}x + \frac{c - c'}{r_1 + \lambda_1}. \quad (4.40)$$

Now we solve for the coefficients  $A_i$ ,  $B_i$ ,  $C_i$  and the optimal values  $v_1$  and  $v_2$ . Using the boundary and smooth pasting conditions for  $P_2(x)$  (see(4.30), (4.33)) at  $v_2$  with  $x \in [v_2, \infty)$ ,

$$\begin{cases} Q_2^* + B_1v_2^{\beta_1} + B_2v_2^{\beta_2} = (1 - \alpha)v_2, \\ \beta_1B_1v_2^{\beta_1-1} + \beta_2B_2v_2^{\beta_2-1} = (1 - \alpha). \end{cases}$$

or

$$\begin{cases} l_1A_1v_2^{\beta_1} + l_2A_2v_2^{\beta_2} = (1 - \alpha)v_2 - Q_2^*, \\ \beta_1l_1A_1v_2^{\beta_1-1} + \beta_2l_2A_2v_2^{\beta_2-1} = (1 - \alpha)v_2. \end{cases} \quad (4.41)$$

Similarly, using the boundary and smooth pasting conditions for  $P_1(x)$  (see (4.29), (4.32)) at  $v_2$  with  $x \in [v_1, v_2]$ ,

$$\begin{cases} Q_1^* + A_1v_2^{\beta_1} + A_2v_2^{\beta_2} = C_1v_2^{\gamma_1} + C_2v_2^{\gamma_2} + \phi(v_2), \\ \beta_1A_1v_2^{\beta_1-1} + \beta_2A_2v_2^{\beta_2-1} = \gamma_1C_1v_2^{\gamma_1-1} + \gamma_2C_2v_2^{\gamma_2-1} + v_2\phi'(v_2). \end{cases} \quad (4.42)$$

Using the boundary and smooth pasting conditions for  $P_1(x)$  (see (4.29), (4.32)) at  $v_1$  with  $x \in [v_1, v_2]$

$$\begin{cases} C_1v_1^{\gamma_1} + C_2v_1^{\gamma_2} + \phi(v_1) = (1 - \alpha)v_1, \\ \gamma_1C_1v_1^{\gamma_1-1} + \gamma_2C_2v_1^{\gamma_2-1} + v_1\phi'(v_1) = (1 - \alpha)v_1. \end{cases} \quad (4.43)$$

Combining these equations, we can obtain the solutions  $v_1$  and  $v_2$ . From (4.43), we have

$$F_1(v_1) := \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} (1 - \alpha)v_1 - \phi(v_1) \\ (1 - \alpha)v_1 - v_1\phi'(v_1) \end{bmatrix} = \begin{bmatrix} C_1v_1^{\gamma_1} \\ C_2v_1^{\gamma_2} \end{bmatrix}. \quad (4.44)$$

From (4.41), we have

$$\begin{bmatrix} l_1 & l_2 \\ \beta_1l_1 & \beta_2l_2 \end{bmatrix}^{-1} \begin{bmatrix} (1 - \alpha)v_2 - Q_2^* \\ (1 - \alpha)v_2 \end{bmatrix} = \begin{bmatrix} A_1v_2^{\beta_1} \\ A_2v_2^{\beta_2} \end{bmatrix}. \quad (4.45)$$

Then, using (4.42) and (4.45), we have

$$\begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_2 - Q_2^* \\ (1-\alpha)v_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} A_1 v_2^{\beta_1} \\ A_2 v_2^{\beta_2} \end{bmatrix} = \begin{bmatrix} C_1 v_2^{\gamma_1} + C_2 v_2^{\gamma_2} + \phi(v_2) - Q_1^* \\ \gamma_1 C_1 v_2^{\gamma_1} + \gamma_2 C_2 v_2^{\gamma_2} + v_2 \phi'(v_2) \end{bmatrix}. \quad (4.46)$$

This equation can be rewritten as

$$\begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_2 - Q_1^* \\ (1-\alpha)v_2 \end{bmatrix} - \begin{bmatrix} \phi(v_2) - Q_1^* \\ v_2 \phi'(v_2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix} \begin{bmatrix} C_1 v_2^{\gamma_1} \\ C_2 v_2^{\gamma_2} \end{bmatrix}.$$

Therefore

$$F_2(v_2) := \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \left( \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_2 - Q_2^* \\ (1-\alpha)v_2 \end{bmatrix} - \begin{bmatrix} \phi(v_2) - Q_1^* \\ v_2 \phi'(v_2) \end{bmatrix} \right) = \begin{bmatrix} C_1 v_2^{\gamma_1} \\ C_2 v_2^{\gamma_2} \end{bmatrix}. \quad (4.47)$$

Hence, from (4.44) and (4.47) we have the equation

$$\begin{bmatrix} v_1^{-\gamma_1} & 0 \\ 0 & v_1^{-\gamma_2} \end{bmatrix} F_1(v_1) = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} F_2(v_2). \quad (4.48)$$

In particular, for  $\phi(x)$  of the form in (4.40), we have from (4.44) and (4.47)

$$F_1(v_1) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_1 - \frac{\lambda_1(1-\alpha)}{r_1+\lambda_1-\mu_1}v_1 - \frac{c-c'}{r_1+\lambda_1} \\ (1-\alpha)v_1 - v_1 \frac{\lambda_1(1-\alpha)}{r_1+\lambda_1-\mu_1} \end{bmatrix}, \quad (4.49)$$

and

$$F_2(v_2) = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \left( \begin{bmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_2 - Q_2^* \\ (1-\alpha)v_2 \end{bmatrix} - \begin{bmatrix} \frac{\lambda_1(1-\alpha)}{r_1+\lambda_1-\mu_1}v_2 + \frac{c-c'}{r_1+\lambda_1} - Q_1^* \\ (1-\alpha)v_2 \frac{\lambda_1}{r_1+\lambda_1-\mu_1} \end{bmatrix} \right). \quad (4.50)$$

Substituting  $F_1(v_1)$  and  $F_2(v_2)$  into (4.48), we can obtain the values for  $v_1$  and  $v_2$ . Now we derive coefficients  $A_i$ ,  $B_i$ , and  $C_i$ . From (4.41), the coefficients  $A_1$  and  $A_2$  are give by

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} l_1 v_2^{\beta_1} & l_2 v_2^{\beta_2} \\ \beta_1 l_1 v_2^{\beta_1} & \beta_2 l_2 v_2^{\beta_2} \end{bmatrix}^{-1} \begin{bmatrix} (1-\alpha)v_2 - Q_2^* \\ (1-\alpha)v_2 \end{bmatrix}, \quad (4.51)$$

and

$$\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} l_1 A_1 \\ l_2 A_2 \end{bmatrix}. \quad (4.52)$$

From (4.43), we have

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} v_1^{\gamma_1} & v_1^{\gamma_2} \\ \gamma_1 v_1^{\gamma_1} & \gamma_2 v_2^{\gamma_2} \end{bmatrix}^{-1} \begin{bmatrix} (1 - \alpha)v_1 - \phi(v_1) \\ (1 - \alpha)v_1 - v_1 \phi'(v_1) \end{bmatrix} \quad (4.53)$$

$$= \begin{bmatrix} v_1^{\gamma_1} & v_1^{\gamma_2} \\ \gamma_1 v_1^{\gamma_1} & \gamma_2 v_1^{\gamma_2} \end{bmatrix}^{-1} \begin{bmatrix} (1 - \alpha)v_1 - \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} v_1 - \frac{c - c'}{r_1 + \lambda_1} \\ (1 - \alpha)v_1 - v_1 \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} \end{bmatrix}. \quad (4.54)$$

With these coefficients, the value functions  $P_1(x)$  and  $P_2(x)$  become

$$P_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} + \frac{(c - c')(\lambda_1 - \lambda_2)}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \phi(x) & \text{if } v_1 \leq x \leq v_2 \\ (1 - \alpha)x & \text{if } x \leq v_1, \end{cases}$$

and

$$P_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} + \frac{(c - c')((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ (1 - \alpha)x & \text{if } x \leq v_2, \end{cases}$$

where

$$\phi(x) = \frac{\lambda_1(1 - \alpha)}{r_1 + \lambda_1 - \mu_1} x + \frac{c - c'}{r_1 + \lambda_1}.$$

□

## Chapter 5

# Credit Spread of Defaultable Corporate Bonds with Regime Switching

### 5.1 Introduction

Credit risk is among the central topics of financial modeling and risk management. There are two conceptual approaches to model credit risk. The first approach, initiated by Merton (1974) and Black and Cox (1976), and referred to as the structural (firm-value based) approach (see Duffie and Singleton (1999)), considers the relationship between a firm's asset value and the firm's default event, and models the evolution of the firm value as an exogenous diffusion process. The default of the firm is triggered when the asset value falls below a default level, which could be time varying. The approach has also been adopted by many researchers (see, e.g., Geske (1977), Ingersoll (1977a, 1977b), Merton (1977), Smith and Warner (1979), Cooper and Mello (1991), Hull and White (1992), Abken (1993), etc.).

An alternative approach, introduced by Jarrow and Turnbull (1995), Jarrow, Lando and Turnbull (1997), Duffie and Singleton (1997, 1999), etc., and known as the reduced-form approach, treats defaults as exogenous events and models the arrivals of defaults using doubly stochastic Poisson point processes. The advantage of the reduced-form approach is that it can generate rich dynamics of the term structure of credit spreads while keeping tractability. However, the reduced-form approach usually does not model explicitly the relationship between firm value and corporate default risk in a structural way. The hazard rate of default in the reduced-form approach is generally modeled as an exogenous process. Therefore, some economic insights or mechanisms are frequently absent with the reduced-form approach. It is usually not clear what economic mechanisms give rise to the default

process, and few theoretical economic insights are given to explain how the dynamics are caused. Duffee (1999) also indicated that the parameters of a reduced-form model could be unstable when one tries to fit the model to observed yield spreads.

In comparison, a reduced-form approach is usually more flexible to fit the observed credit spreads, while a structural approach often generates more economic insights and conceptual mechanisms on default behavior.

The original models considered by Merton (1974), Black and Cox (1976) and Brennan and Schwartz (1978) etc. model the evolution of firm value as a diffusion process. This assumption results in unrealistic empirical implications. For example, if a firm cannot default unexpectedly and, unless it is not currently in financial distress, its probability of defaulting on very short-term debt is zero and therefore, its short-term debt should have zero credit spreads and its term structure of credit spreads should slope upward at the short end. Empirical evidence indicates that credit spreads on typical short term bonds are essentially greater than zero. Moreover, Fons (1994) and Sarig and Warga (1989) show that the yield spread curves of certain kinds of bonds are even downward sloping. The empirical analysis in Jones, Mason, and Rosenfeld (1984) indicate that the credit spreads on corporate bonds are generally too high to be matched by this approach. Other empirical contributions that actually implement structural models include Delianedis and Geske (1999), Eom et al. (2004), Ericsson and Reneby (2002, 2005), and Huang and Huang (2012).

Researchers have proposed extensions of the structural approach. For example, the structural approach is later extended to allow for stochastic interest rates in Longstaff and Schwartz (1995), Kim et al. (1993), Cathart and El-Jahel (1998), Briys and de Varenne (1997), Nielsen et al. (1993), Saá-Requejo and Santa-Calra (1999), etc.; in Zhou (2001), the author models the evolution of the firm value as a jump-diffusion process, as introduced in Merton (1973).

On the other hand, there have been a large amounts of empirical evidence suggesting that



the aggregate economy is characterized by periodic shifts between distinct business cycles (see, e.g., Hamilton (1989)). In addition to the empirical evidence, there are also economic reasons to believe that regime switching is important to understand the dynamics of the credit spread curve. Empirically, the expansion and contraction periods in business cycles have potentially sizable effects on macroeconomic fundamentals, such as inflationary expectations, monetary policy, and nominal interest rates. Also, these regime shifts change corporations' growth prospects, and affect corporate earnings, as well as investment or default decisions. On economic grounds, such switchings in regimes give rise to the possibility of significant impacts on firms' profitability, riskiness of assets and investments, and the likelihood of debt defaults. Therefore, the regime switchings can affect firms' capital accumulation, and policy choices, as well as investment or default decisions, and consequently affect the credit risk associated with these firms, as well as the credit spreads of the credit risk bearing securities issued by the firms. In particular, Giesecke et al. (2011) conducted empirical analyses on the corporate bond market over the past 150 years. Their empirical data suggests that there exist three regimes, and they also obtain the transitional probability.

In a typical structural model, e.g., Merton (1974) and Black and Cox (1976), Leland (1994), etc., credit risk is analyzed by viewing equities as options written on real assets. Because option values depend on the riskiness of the underlying assets, volatility is an important determinant of the default risk. In previous contributions, e.g., Merton (1974), Black and Cox (1976), Longstaff and Schwartz (1995), Leland (1994), etc., models typically postulate that this parameter is a fixed constant. It is natural to imagine, as interpreted above, that as volatility changes over the business cycle, so does the credit (default) risk. Indeed, Huang and Zhou (2008) conducted empirical analyses to test these models, and their empirical tests strongly reject the models.

In this paper, we develop a framework to model credit risk when the dynamics of the state variables are subject to regime shifts at random times. Following Hamilton (1989),

we think of these regime switchings as “episodes across which the behavior of the series is markedly different”. We construct two models that allow for the regime switching risk, which extend earlier contributions from Merton (1974), Black and Cox (1976), etc. Moreover, we are able to obtain closed-form explicit solutions for the credit spread of defaultable debts in these models, where regime switchings are allowed.

The paper is organized as follows. Section 2 presents a basic model where default can occur only at maturity. This is a model extending the analysis in Merton (1974). In section 3, we consider a model where endogenous default is allowed, and default can occur anytime prior to or at maturity. This model extends the classic work in Black and Cox (1976). For both models, we are able to obtain closed-form analytical solutions for the value and credit spread of defaultable bonds. In section 4 we present some concluding remarks.

## 5.2 Model with Default Only at Maturity

Suppose a firm issues two types of claims, equity and defaultable debts. For simplicity, we suppose the debt is a zero-coupon bond with par value  $D$  and maturity time  $T$ . At the maturity time  $T$ , the payoffs to the debt and the equity are denoted by  $D(T, T)$  and  $E_T$ , respectively.

The conceptual foundations for default occurrence were laid by Merton (1970, 1974) and Black and Scholes (1973). Following their assumption, we suppose that possible default occurs only at the maturity date of the debt, provided that the issuer’s assets are less than the face value of maturing debt at that time. Suppose the value of the firm’s assets at maturity  $T$  is  $V_T$ . At the maturity of the debt, the firm pays to the debt holder precisely the face value of the debt,  $D$ , if the asset value  $V_T$  is higher than the par value of the debt. On the other hand, if assets are worth less than  $D$  at maturity, bond holders take over the assets and receive a “recovery value”  $V_T$  instead of the promised payment  $D$ . Thus, the payoffs

$D(T, T)$  and  $E_T$  are given by

$$D(T, T) = \min(D, V_T) = D - \max(D - V_T, 0),$$

$$E_T = \max(V_T - D, 0) = (V_T - D)^+.$$

The question we consider is how the debt and equity are valued prior to the maturity time  $T$ . We note from the structure of the payoffs that  $V_T = D(T, T) + E_T$ , and the debt can be viewed as a riskless bond less a put option, and equity can be viewed as a call option on the firm's assets.

We suppose the state of the economy is described by a finite state continuous-time Markov chain  $X = \{X_t, t \geq 0\}$ . In our model, the states of the Markov Chain represent different economic environments. For example, there could be just two states for  $X$  representing “good” and “bad” economic regimes. The switching of the states of the economy can be attributed to structural changes in macroeconomic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities.

For notational convenience, the state space of  $X$  can be taken to be, without loss of generality, the set of unit vectors  $\{e_1, e_2, \dots, e_N\}$ , where  $e_i = (0, \dots, 0, 1, 0, \dots, 0)' \in \mathbb{R}^N$  with the “1” in the  $i$ th entry and where the superscript  $'$  represents the transpose of a row vector. This is called the *canonical representation* of the state space of the Markov chain  $X$ . Suppose the process  $X$  is homogeneous in time. Let  $A$  denote the constant rate matrix of the Markov chain  $X$ . Then, using the canonical representation of the state space, the dynamics of the Markov chain  $X$  have the following *semi-martingale representation* (see Elliott et al. (1995))

$$dX_t = AX_t dt + dM_t, \tag{5.1}$$

where  $(M_t)_{t \geq 0}$  is an  $\mathbb{R}^N$ -valued martingale with respect to the natural filtration generated by  $X$ .

We suppose the short (instantaneous) interest rate  $r$  depends on the state  $X_t$  of the economy, so that

$$r_t = r(X_t) = \langle r, X_t \rangle,$$

where  $r = (r_1, r_2, \dots, r_N)' \in \mathbb{R}^N$ , and each  $r_i$  is a constant,  $1 \leq i \leq N$ .

Suppose  $W$  is a standard Brownian motion on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Suppose the dynamics of the firm's assets are driven by a regime switching geometric Brownian motion

$$\frac{dV_t}{V_t} = (r_t - a_t)dt + \sigma_t dW_t,$$

where  $a_t = a(X_t)$  is the payout rate at time  $t$ , which is also switching with the Markov chain  $X$ . Thus, there exists a vector  $a := (a_1, \dots, a_N)$  such that  $a_t = \langle a, X - t \rangle$ . Let  $\mu_t := r_t - a_t$ .

Then

$$\frac{dV_t}{V_t} = \mu_t dt + \sigma_t dW_t.$$

Here, the drift  $\mu_t = \mu(X_t)$  and volatility  $\sigma_t = \sigma(X_t)$  both depends on the state  $X_t$  of the economy. That is, there are vectors  $\mu = (\mu_1, \dots, \mu_N)' = (r_1 - a_1, \dots, r_N - a_N)'$  and  $\sigma = (\sigma_1, \dots, \sigma_N)'$  such that  $\mu_t = \langle \mu, X_t \rangle$  and  $\sigma_t = \langle \sigma, X_t \rangle$ .

Let  $C(t, T, V_0, X_0)$  denote the arbitrage-free price at time  $t \in [0, T]$  of a European call option with strike price  $D$  and maturity time  $T$ . Then

$$C(t, T, V_t, X_t) = \mathbb{E}_{\mathbb{Q}} \left[ e^{-\int_t^T r_u du} (V_T - D)^+ | \mathcal{F}_t \right],$$

where  $\mathcal{F}_t = \sigma\{W_u, X_u : 0 \leq u \leq t\}$ .

As noted before, the equity can be viewed as a call option with strike  $D$  on the firm's assets, and the debt can be viewed as a riskless bond less a put option. Therefore, the values of the equity and the debt at time  $t$  are

$$S(t, T) = C(t, T, V_t, X_t),$$

$$D(t, T) = De^{-\int_t^T r_u du} - P(t, T, V_t, X_t).$$

Here,  $P(t, T, V_t, X_t)$  denotes the arbitrage-free price for the European put option, which can be obtained by the put–call parity, i.e.,

$$C(t, T, V_t, X_t) - P(t, T, V_t, X_t) = V_t - De^{-\int_t^T r_u du}.$$

Using the put–call parity, we have an alternative formula for the value of the defaultable debt at time  $t$ :

$$D(t, T) = V_t - C(t, T, V_t, X_t).$$

To obtain the value of the defaultable bond, we need only find the price of the European call or put, with strike price  $D$ . It is shown in Buffington and Elliott (2002) that, the price  $C$  of the European call satisfies the coupled PDE

$$\frac{\partial C}{\partial t} + \mu_t V_t \frac{\partial C}{\partial V} + \frac{1}{2}(\sigma_t^2 V_t)^2 \frac{\partial^2 C}{\partial V^2} - r_t C + \langle \mathbf{C}, AX_t \rangle = 0,$$

with terminal condition

$$C(T, T, V_T, X_T) = (V_T - D)^+.$$

Here  $\mathbf{C} = (C(t, T, V_t, e_1), C(t, T, V_t, e_2), \dots, C(t, T, V_t, e_K))$ . Writing  $C_i = C(t, T, V_t, e_i)$  so that  $\mathbf{C} = (C_1, C_2, \dots, C_K)$ , we see that each component  $C_i$  of  $\mathbf{C}$  satisfies the PDE

$$\frac{\partial C_i}{\partial t} + \mu_t V_t \frac{\partial C_i}{\partial V} + \frac{1}{2}(\sigma_t^2 V_t)^2 \frac{\partial^2 C_i}{\partial V^2} - r_t C + \langle \mathbf{C}, Ae_i \rangle = 0,$$

with terminal condition

$$C_i(T, T, V_T, e_i) = (V_T - D)^+.$$

Let  $P_T = \int_t^T \langle r_u, X_u \rangle du$ ,  $U_T = \int_t^T \langle \sigma_u, X_u \rangle^2 du$ . For  $1 \leq i \leq N$ , let  $T_i = \int_0^T \langle e_i, X_u \rangle du$  denote the amount of time  $X$  has spent in state  $e_i$  up to time  $T$ . Then

$$\begin{aligned} P_T &= \sum_{i=1}^N r_i T_i = \sum_{i=1}^{N-1} (r_i - r_N) T_i + r_N T, \\ U_T &= \sum_{i=1}^N \sigma_i^2 T_i = \sum_{i=1}^{N-1} (\sigma_i^2 - \sigma_N^2) T_i + r_N T. \end{aligned}$$

If we knew the trajectory of  $X$  over the time horizon  $[, 0, T]$ , we would know the value of  $P_T$  and  $U_T$ . Thus, conditional on the trajectory of  $X$ , we know the value of a European call, which is given by

$$C(0, T, V_0, P_T, U_T) = V_t N(d_1) - e^{P_T} D N(d_2),$$

where

$$\begin{aligned} d_1 &= U_T^{-1/2} (\log(V_t/D) + P_T + U_T/2), \\ d_2 &= d_1 - U_T^{1/2} \end{aligned}$$

To obtain the value of the European call in a regime switching economy, a second expectation over the variables  $P_T$  and  $U_T$  must be taken. The distributions of  $P_T$  and  $U_T$ , by their definition, are determined by the joint probability distributions of the amount of time  $X$  has spent in state  $e_i$  up to time  $T$ . Therefore, it suffices to find the joint density function  $\phi(T_1, T_2, \dots, T_{N-1})$  for the random vector  $(T_1, \dots, T_{N-1})$ . The Fourier transform of the density  $\phi(T_1, T_2, \dots, T_{N-1})$  for  $(T_1, \dots, T_{N-1})$  has been derived in Buffington and Elliott (2002).

**Lemma 5.2.1** *Suppose  $\phi(T_1, T_2, \dots, T_{N-1})$  is the density function of  $(T_1, T_2, \dots, T_{N-1})$ . Then the Fourier transform of  $\phi(T_1, T_2, \dots, T_{N-1})$  is*

$$\mathcal{F}[\phi(T_1, T_2, \dots, T_{N-1})](\theta_1, \dots, \theta_{N-1}) = \langle \exp((A + i \text{diag } \boldsymbol{\theta})T), \mathbf{1} \rangle, \quad (5.2)$$

where  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_{N-1}, 0) \in \mathbb{R}^N$ .

*Proof.* See Buffington and Elliott (2002). For completeness, we include a proof in the Appendix. □

In next theorem, we obtain the value of the European call in our regime switching economy, using a Fourier inverse-transform.

**Theorem 5.2.2** *The value of the European call is*

$$C(0, T, V_0, X_0) = \int_{[0,1]^{N-1}} C(0, T, V_0, P_T, U_T) \phi(T_1, T_2, \dots, T_{N-1}) d(T_1, T_2, \dots, T_{N-1}),$$

where  $\phi(T_1, T_2, \dots, T_{N-1})$  is the density function of  $(T_1, T_2, \dots, T_{N-1})$ , given by

$$\phi(T_1, T_2, \dots, T_{N-1}) = \frac{1}{(2\pi)^{N-1}} \int_{[0,1]^{N-1}} \langle \exp((A + i \operatorname{diag} \boldsymbol{\theta})T), \mathbf{1} \rangle e^{-i\boldsymbol{\theta}T} d(\theta_1, \theta_2, \dots, \theta_{N-1}).$$

*Proof.* From Lemma 5.2.1, we see that the Fourier transform of  $\phi(T_1, T_2, \dots, T_{N-1})$  is

$$\mathcal{F}[\phi](\theta_1, \theta_2, \dots, \theta_{N-1}) = \langle \exp((A + i \operatorname{diag} \boldsymbol{\theta})T), \mathbf{1} \rangle. \quad (5.3)$$

Taking Fourier inverse transform, we see that

$$\phi(T_1, T_2, \dots, T_{N-1}) = \frac{1}{(2\pi)^{N-1}} \int_{[0,1]^{N-1}} \langle \exp((A + i \operatorname{diag} \boldsymbol{\theta})T), \mathbf{1} \rangle e^{-i\boldsymbol{\theta}T} d(\theta_1, \theta_2, \dots, \theta_{N-1}).$$

The value of the European option is the expectation of  $C(0, T, V_0, P_T, U_T)$  over all trajectories of  $X$  on  $[0, T]$ . That is the expectation of  $C(0, T, V_0, P_T, U_T)$  over the joint distribution of  $(T_1, T_2, \dots, T_{N-1})$ . Thus

$$C(0, T, V_0, X_0) = \int_{[0,1]^{N-1}} C(0, T, V_0, P_T, U_T) \phi(T_1, T_2, \dots, T_{N-1}) d(T_1, T_2, \dots, T_{N-1}),$$

where  $\phi(T_1, T_2, \dots, T_{N-1})$  has obtained with the Fourier inverse transform.  $\square$

In particular, if  $X$  has two states, i.e.,  $X = \{e_1, e_2\}$ , where  $e_1 = (1, 0)'$ , and  $e_2 = (0, 1)'$ , the Fourier transform (characteristic function) of  $\phi(T_1)$  can be obtained in a more explicit and compact form. For this two-state economy, Buffington and Elliott (2002) derived a formula with an error. We derive a corrected closed-form formula, and provide a proof.

**Lemma 5.2.3** *When  $N = 2$ , the Fourier transform of  $\phi(T_1)$  is*

$$\mathcal{F}[\phi(T_1)](\theta_1) = e^{a_{22}t} \left( c_1 e^{y_1 t} \left( 1 - \frac{a_{11}}{y_1} \right) + c_2 e^{y_2 t} \left( 1 - \frac{a_{11}}{y_2} \right) + \langle e_2, X_0 \rangle + a_{11} \left( \frac{c_1}{y_1} + \frac{c_2}{y_2} \right) \right),$$

where  $y_1, y_2, c_1$ , and  $c_2$  are given in equations (5.7), (5.8), and (5.9).

*Proof.* See Appendix. □

**Theorem 5.2.4** *If  $X$  has two states, the value of the European call is*

$$\begin{aligned} C(0, T, V_0, X_0) &= \int_0^T C(0, T, V_0, P_T, U_T) \phi(T_1) dT_1 \\ &= \int_0^T C(0, T, V_0, (r_1 - r_2)T_1 + r_2T, (\sigma_1^2 - \sigma_2^2)T_1 + \sigma_2^2T) \phi(T_1) dT_1, \end{aligned}$$

where  $\phi(T_1)$  is the density function of  $T_1$ , and is given by

$$\phi(T_1) = \frac{1}{2\pi} \int_0^T e^{a_{22}t} \left( c_1 e^{y_1 t} \left( 1 - \frac{a_{11}}{y_1} \right) + c_2 e^{y_2 t} \left( 1 - \frac{a_{11}}{y_2} \right) + \langle e_2, X_0 \rangle + a_{11} \left( \frac{c_1}{y_1} + \frac{c_2}{y_2} \right) \right) e^{-i\theta_1 T_1} d\theta_1.$$

*Proof.* Similar to the proof of Theorem 5.2.2. □

**Theorem 5.2.5** *Suppose the dynamics of the firm's assets are driven by a regime switching geometric Brownian motion*

$$\frac{dV_t}{V_t} = \mu_t dt + \sigma_t dW_t.$$

Then the value of the defaultable corporate debt is

$$D(0, T; X) = V_0 - C(0, T, V_0, X_0).$$

Here  $C(0, T, V_0, X_0)$  is given in Theorems 5.2.2 and 5.2.4.

Since corporate debts typically have promised cash flows which resemble those of treasury bonds, it is more appropriate to consider yields rather than prices. The *yield* at time  $t$  of a bond with maturity  $T$  is the quantity  $y(t, T; X)$  such that

$$D(t, T; X) e^{y(t, T; X)(T-t)} = D,$$

i.e., the yield is given by

$$y(t, T; X) = \frac{1}{T-t} \log \left( \frac{D}{D(t, T; X)} \right).$$

To compare the difference between the yields of a corporate bond and a treasury bond, we consider the credit spread, or yield spread, of the corporate bond. The *credit spread*, or *yield*



*spread*, of a defaultable corporate bond is the difference between the yields of the corporate bond and the corresponding treasury bond, i.e.,

$$s(t, T; X) = y(t, T; X) - r_t.$$

When  $t = 0$ , we write  $s(T)$  for  $s(0, T)$ . The risk structure of interest rates is obtained when one views  $s(T)$  as a function of  $T$ .

**Proposition 5.2.6** *The credit spread of the defaultable corporate bond is*

$$s(T; X) = y(0, T; X) - r = \frac{1}{T} \log \left( \frac{D}{D(0, T; X)} \right) - r_t$$

where  $D(0, T; X) = V_0 - C(0, T, V_0, X_0)$  is given in Theorems 5.2.2 and 5.2.4.

Empirical evidence indicates that credit spreads on typical short term bonds are essentially greater than zero. The empirical analysis in Jones, Mason, and Rosenfeld (1984) indicate that the credit spreads on corporate bonds are generally too high to be matched by the previous models.

With regime switching risk, additional risk premium for regime shift is required. And so models can naturally generate higher credit spread. The models with regime switching are potentially able to better match the actual, empirical credit spread.

### 5.3 Model with First-Passage-Time Default

The model in Section 2 allows credit event to occur only at debt's maturity. In this section we consider a first-passage-time model with regime switching. Thus default may occur at any time before or on the bond's maturity date  $T$ . This model accounts for the observed feature that the default may well occur not only at the debt's maturity, but also prior to maturity. More precisely, we consider an extension of the Black-Cox model with regime

switching. The value of the firm's assets follows dynamics<sup>1</sup>

$$\begin{aligned}\frac{dV_t}{V_t} &= (r(X_t) - a(X_t))dt + \sigma(X_t)dW_t \\ &= \langle r - a, X_t \rangle dt + \langle \sigma, X_t \rangle dW_t.\end{aligned}$$

Here  $r(X_t)$ ,  $a(X_t)$ , and  $\sigma(X_t)$  are deterministic functions of  $X_t$  representing the default-free interest rate, the payout ratio, and the volatility of the assets at time  $t$ , respectively. As before, we have  $r(X_t) = \langle r, X_t \rangle$  for  $r = (r_1, \dots, r_N)$ ,  $a(X_t) = \langle a, X_t \rangle$  for  $a = (a_1, \dots, a_N)$ , and  $\sigma(X_t) = \langle \sigma, X_t \rangle$  for  $\sigma = (\sigma_1, \dots, \sigma_N)$ .

### 5.3.1 Coupled PDE for General Cash Flow Stream

We first derive the coupled PDE for the value of the coupon bond. We can obtain the following theorem for a general cash flow stream. Let  $S$  be a process, which models the dynamics of an underlying risky asset, with

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t,$$

where  $\mu_t = \mu(X_t) = \langle \mu, X \rangle$ ,  $\sigma_t = \sigma(X_t) = \langle \sigma, X \rangle$ . We further assume that  $r_t = r(X_t) = \langle r, X \rangle$ , where  $r := (r_1, \dots, r_N)$ .

Consider a contingent claim which has a continuous dividend stream  $g(t, S_t, X_t)$  at time  $t$ , and a terminal payment  $f(S_T, X_T)$  at maturity  $T$ . The value of the contingent claim at time  $t$  is

$$P_i(t, S) = \mathbb{E} \left[ \int_t^T e^{-\int_t^u r_s ds} g(u, S_u) du + e^{-\int_t^T r_u du} f(S_T) \mid S_t = S, X_t = e_i \right].$$

We can write this in an equivalent form

$$P(t, S, X) = \mathbb{E} \left[ \int_t^T e^{-\int_t^u r_v dv} g(u, S_u, X_u) du + e^{-\int_t^T r_u du} f(S_T, X_T) \mid S_t = S, X_t = X \right],$$

Here  $X = X_t$  takes on value in  $\{e_1, \dots, e_N\}$ .

---

<sup>1</sup>As in Chapter 4 as well as the previous contributions, e.g. Leland (1994) and Leland and Toft (1996), we assume, without loss of generality, the dynamics are in the risk-neutral measure.

**Theorem 5.3.1** *Under some technical conditions, the function  $P(t, S, X) := P(t, S_t, X_t)$  is the solution of the coupled PDE*

$$P_t + \frac{1}{2}\sigma^2 S^2 P_{SS} + \mu S P_S - rP + g + \langle P, AX \rangle = 0,$$

$$P(T, x, X) = f(x, X_T).$$

*Proof.* Let  $P(t, S, X) := P(t, S_t, X_t)$ . Suppose  $P(t, S, X)$  satisfies the coupled PDE. We consider  $e^{-\int_t^s r_u du} P(s, S_s, X_s)$ . Using Itô's rule, we have

$$d\left(e^{-\int_t^s r_u du} P(s, S_s, X_s)\right) = e^{-\int_t^s r_u du} \left(P_t + \frac{1}{2}\sigma^2 S^2 P_{SS} + \mu S P_S - rP + \langle P, AX \rangle\right) dt$$

$$+ [\dots] dW_t + e^{-\int_0^t r_u du} \langle P, dM_t \rangle.$$

Integrating from  $t$  to  $T$  (letting variable  $s$  range from  $t$  to  $T$ ), we have

$$e^{-\int_t^T r_u du} P(T, S_T, X_T) - e^{-\int_t^t r_u du} P(t, S_t, X_t) = \int_t^T e^{-\int_t^s r_u du} \left(P_t + \frac{1}{2}\sigma^2 S^2 P_{SS} + \mu S P_S - rP + \langle P, AX \rangle\right) dt$$

$$+ [\dots] dW_t + e^{-\int_0^t r_u du} \langle P, dM_t \rangle$$

$$= \int_t^T e^{-\int_t^v r_u du} (-g(v, S_v, X_v)) dv + \int_t^T [\dots] dW_t + \int_t^T e^{-\int_0^t r_u du} \langle P, dM_t \rangle.$$

(from coupled PDE)

Using the boundary condition in the coupled PDE, we see that

$$e^{-\int_t^T r_u du} f(S_T, X_T) + \int_t^T e^{-\int_t^s r_u du} g(v, S_v, X_v) dv - [\dots] dW_t = P(t, S_t, X_t)$$

Under technical regularity conditions, the conditional expectation  $\mathbb{E}[\dots dW_t | \mathcal{F}_t]$  of the  $dW_t$ -term vanishes. Thus, taking conditional expectation  $\mathbb{E}[\cdot | \mathcal{F}_t]$ , we see that,

$$P(t, S, X) = \mathbb{E} \left[ e^{-\int_t^T r_u du} f(S_T, X_T) + \int_t^T e^{-\int_t^s r_u du} g(v, S_v, X_v) dv | \mathcal{F}_t \right]$$

$$= \mathbb{E} \left[ e^{-\int_t^T r_u du} f(S_T, X_T) + \int_t^T e^{-\int_t^s r_u du} g(v, S_v, X_v) dv | S_t = S, X_t = X \right].$$

This is the definition of  $P(t, S, X)$ . □

### 5.3.2 Defaultable Coupon Bond

We consider the safety covenants in the firm's indenture provisions. The safety covenants provide the firm's debtholders with the right to force the firm to bankruptcy or reorganization if the firm defaults. Following the idea proposed in Black and Cox (1976), we suppose the performance standard is represented in terms of a time-dependent deterministic barrier  $v_B(t) = K(X_t)e^{-\int_t^T \gamma_u du}$ ,  $t \in [0, T)$ , for some deterministic function  $K(X_t)$  of  $X_t$  and a time-varying deterministic function  $\gamma_t = \gamma(t)$ . Note that  $\gamma_t = \gamma(t)$  is just a **deterministic function** of the time index  $t$ , and not a function of  $X_t$ , i.e.,  $\gamma_t$  is not stochastic. We postulate that the firm defaults as soon as the value of the firm's asset crosses this lower threshold and, in the event of default, debt holders take over the firm. If the firm has not defaulted at time  $t \in [0, T)$ , then default happens at debt's maturity time  $T$  depending on whether or not  $V_T < D$ , the par of the debt. We suppose the debt is a coupon bond, paying continuous coupon at rate  $c$  at every instant of time  $t$ .

From Theorem 5.3.1, we see that the coupled PDEs for the pricing function  $P(t, x, X)$  of a consol bond is

$$\begin{aligned} P_t + \frac{1}{2}x^2\sigma^2P_{xx} + (r_t - \kappa_t)xP_x - r_tP + c + \langle P, AX \rangle &= 0, \\ P(t, v_B(t), X) &= v_B(t) = K(X_t)e^{-\int_t^T \gamma_u du}, \\ P(T, x, X) &= f(x, X_T). \end{aligned}$$

We further assume that  $K(X_t) \leq D$ , as the payoff to the bondholders at the default time never exceeds the face value of debt, discounted at a risk-free rate.

Alternatively, we can define a *fictitious* asset  $S$  so that the default barrier for  $S$  is  $\langle K, X_t \rangle$ .

Define

$$S_t = e^{-\int_0^t \gamma_u du} V_t.$$

**Lemma 5.3.2** *The dynamics for  $S$  is,  $S_0 = V_0$ , and*

$$\frac{dS_t}{S_t} = (r(X_t) - a(X_t) - \gamma(t))dt + \sigma(X_t)dW_t$$

$$= \langle r - a - \gamma, X_t \rangle dt + \langle \sigma, X_t \rangle dW_t.$$

*Proof.* Note that

$$V_t = V_0 \exp \left( \int_0^t (r_u - a_u - \frac{1}{2} \sigma_u^2) du + \int_0^t \sigma_u dW_u \right).$$

Then

$$S_t = V_0 \exp \left( \int_0^t (r_u - a_u - \frac{1}{2} \sigma_u^2 - \gamma_u) du + \int_0^t \sigma_u dW_u \right).$$

Thus, using the Itô Lemma,

$$\begin{aligned} \frac{dS_t}{S_t} &= (r_t - a_t - \gamma_t) dt + \sigma(X_t) dW_t \\ &= \langle r - a - \gamma, X_t \rangle dt + \langle \sigma, X_t \rangle dW_t, \end{aligned}$$

and  $S_0 = V_0$ . □

Write  $\mu_t := r_t - a_t - \gamma_t$ . Then the above dynamics obtained in Lemma 5.3.2 are

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu_t dt + \sigma_t dW_t \\ &= \langle \mu, X_t \rangle dt + \langle \sigma, X_t \rangle dW_t. \end{aligned}$$

**Proposition 5.3.3** *The first time when the value of the firm's asset  $P_t$  crosses the lower threshold  $v_B(t) = K(X_t)e^{-\int_t^T \gamma_u du}$  is the same time when the value  $S_t$  crosses the barrier  $\bar{v}(X_t) := K(X_t)e^{-\int_0^T \gamma_u du}$ .*

*Proof.* The default time  $\tau$  is determined by

$$\begin{aligned} \tau &= \inf \left\{ 0 \leq t \leq T : V_t \leq K_t e^{-\int_t^T \gamma(u) du} \right\} \\ &= \inf \left\{ 0 \leq t \leq T : V_0 \exp \left( \int_0^t (r_u - a_u - \frac{1}{2} \sigma_u^2) dt + \int_0^t \sigma_u dW_u \right) \leq K(X_t) e^{-\int_t^T \gamma(u) du} \right\} \\ &= \inf \left\{ 0 \leq t \leq T : \log V_0 + \left( \int_0^t (r_u - a_u - \frac{1}{2} \sigma_u^2) du + \int_0^t \sigma_u dW_u \right) \leq \log (\langle K, X_t \rangle) - \int_t^T \gamma_u du \right\} \\ &= \inf \left\{ 0 \leq t \leq T : \log V_0 + \left( \int_0^t (r_u - a_u - \frac{1}{2} \sigma_u^2 - \gamma_u) du + \int_0^t \sigma_u dW_u \right) \leq \log (\langle K, X_t \rangle) - \int_0^T \gamma_u du \right\} \end{aligned}$$

$$\begin{aligned}
&= \inf \left\{ 0 \leq t \leq T : V_0 \exp \left( \int_0^t (r_u - a_u - \frac{1}{2} \sigma_u^2 - \gamma_u) du + \int_0^t \sigma_u dW_u \right) \leq K_t e^{-\int_0^T \gamma_u du} \right\} \\
&= \inf \{ 0 \leq t \leq T : S_t \leq K_t \}.
\end{aligned}$$

□

*Remark.* We note that  $\bar{v}(X_t) = K(X_t) e^{-\int_0^T \gamma_u du}$  is a finite-valued function. We write  $v_i := \bar{v}(e_i) = K_i e^{-\int_0^T \gamma_u du}$  for  $1 \leq i \leq N$ . Thus, each  $v_i$  is a constant across the time horizon  $[0, T]$ .

Using Theorem 5.3.1, we have the coupled PDE for the value of the coupon bond  $P(t, x, X) = P(t, S_t, X_t)$ , where  $S_t = x$ ,  $X$  takes on value in  $\{e_1, \dots, e_N\}$ :

$$\begin{aligned}
P_t + \frac{1}{2} x^2 \sigma^2 P_{xx} + (r_t - a_t - \gamma_t) x P_x - r_t P + c + \langle P, AX \rangle &= 0, \\
P(t, v_i, X) &= K_i e^{-\int_t^T \gamma_u du} = v_i e^{\int_0^t \gamma_u du}, \\
P(T, x, X) &= f(x, X_T).
\end{aligned}$$

To find an explicit solution to this problem, we consider a perpetual bond, so that the value function is independent of time  $t$ .

### 5.3.3 Perpetual Bond

In this section, we consider consol (perpetual) bonds, and derive closed-form analytical solution for the value of a defaultable bond.

In reality, a perpetual bond is not as unreasonable as people might think. Although debt securities generally have a finite maturity and, therefore, a time-dependent cash flow stream and values, a perpetual bond is of importance in debt pricing theory for a number of reasons. Firstly, perpetual bonds can be regarded as an approximation of debt with long maturity. If debt has a sufficiently long maturity, then the return of principle has little value and, thus, can be effectively ignored in financial modeling. For example, for a 30-year debt, when interest rate  $r$  equals 10%, the final repayment of principle  $D$  is less than 6% of debt value; when interest rate  $r$  increases to 15%, the final repayment of principle  $D$  is down to

only 1.5% of debt value. Nowadays, it is not unusual that firms issue long terms bonds, e.g., 50-year bonds. Recently, Disney has even issued 100-year debt. Secondly, perpetual debt is not new. For example, Modigliana and Miller (1958), Merton (1974), Black and Cox (1976), and Leland (1994) all consider corporate debt with infinite maturity. Moreover, the Bank of England has since 1752 issued consol bonds, which promise a fixed coupon with no final maturity. Also, similarly to a perpetual debt in principle, preferred equity typically pays a fixed dividend without a maturity time.

For a perpetual bond, i.e.,  $T \rightarrow \infty$ , we suppose the default barrier is a finite-valued process depending on  $X_t$ . More precisely,  $v_B(t) = \bar{v}(X_t) = \langle \bar{v}, X_t \rangle$ , where  $\bar{v} = (v_1, \dots, v_N)$ , and  $v_i$ 's are constants over the time horizon  $[0, \infty)^2$ .

Since the value of the perpetual bond is now independent of time index  $t$ , the coupled ODE for the pricing function  $P(x)$  of a consol bond is

$$\frac{1}{2}x^2\sigma^2P''(x) + \mu xP'(x) - rP(x) + c + \langle P, AX \rangle = 0, \quad (5.4)$$

where  $\mu = r_t - \kappa_t - \gamma_t$ , subject to the boundary condition

$$P(v_i) = \min(v_i, c/r),$$

$$\lim_{x \rightarrow \infty} P(x) = c/r.$$

For simplicity, we examine the case where  $X$  has two states, i.e., the economy is shifting

---

<sup>2</sup>When  $T \rightarrow \infty$ , the assumption that  $v_B(t) = K(X_t)e^{-\int_0^\infty \gamma_u du}$  in the previous section does not make much sense here, as frequently  $\int_0^T \gamma_u du \rightarrow \infty$  as  $T \rightarrow \infty$  (The simplest example is when  $\gamma_u$  is a constant, which is the assumption made in Black and Cox (1976) and other subsequent contributions). This implies that the default barrier  $K(X_t)e^{-\int_0^\infty \gamma_u du} = 0$  for any  $t \in [0, \infty)$ , which is unrealistic. For a more interesting question, we consider instead the more realistic assumption that default barrier is a finite-valued process  $v_B(t) = \bar{v}(X_t)$  depending on  $X_t$ . If  $X_t = e_i$ , then  $\bar{v}(e_i) = v_i$ , for  $1 \leq i \leq N$ , where  $v_i$ 's are constants over time horizon  $[0, \infty)$ . Note that in the degenerate case where  $N = 1$ , i.e., no regime switching,  $\bar{v}$  is a constant (scalar) over time horizon  $[0, \infty)$ . This is consistent with the conventional assumption made in previous contributions in, e.g., Black and Cox(1976), Longstaff and Schwartz (1995), Leland (1994), and leland (1996), etc.

between two regimes. Then, the rate matrix  $A$  for the Markov chain  $X$  is of form<sup>3</sup>

$$A = \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{bmatrix}' = \begin{bmatrix} -\lambda_1 & \lambda_2 \\ \lambda_1 & -\lambda_2 \end{bmatrix}.$$

We shall obtain a *closed-form* analytical expression for the value of corporation's debt.

**Theorem 5.3.4** *The value functions of the bond  $P_1(x) := P(x, e_1)$  and  $P_2(x) := P(x, e_2)$*

*are*

$$P_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} + \frac{c(\lambda_1 - \lambda_2)}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \frac{\lambda_1}{r + \lambda_1 - \mu_1} x + \frac{c}{r + \lambda_1} & \text{if } v_1 \leq x \leq v_2 \\ x & \text{if } x \leq v_1, \end{cases}$$

*and*

$$P_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} + \frac{c((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ x & \text{if } x \leq v_2, \end{cases}$$

where  $\beta_1, \beta_2$  are the two negative roots of equation (5.17), and  $\gamma_1, \gamma_2$  are the two roots of equation (5.24). The coefficients  $A_i, B_i$  and  $C_i$  are given in equations (5.34), (5.19), (5.20), and (5.29).

*Proof.* See Appendix. □

*Remark.* Our essential idea and approach can be extended to obtain closed-form solutions for  $N \geq 3$ , i.e., when  $X$  has  $n \geq 3$  states. For example, with similar method, for  $N = 3$ , we can derive systems of equations from the characteristic functions, and obtain closed-form solution. However, when  $N$  becomes large, the number of equations that one derives from the characteristic functions grows rapidly, and thus the procedure becomes more involved and complicated. When  $N$  is reasonably large, e.g.,  $N \geq 5$ , it is perhaps better to use a software package to handle the systems of equations.

---

<sup>3</sup>Note that in our setting  $A$  is the transpose of the usual generator matrix of a Markov chain.



## 5.4 Concluding Remarks

We investigate credit risk and the credit spread of a corporate defaultable bond when the dynamics of the assets (both growth rate and volatility) shift between different states of the economy.

We consider two models. The first model considers the case when the firm's default may occur only at maturity with regime switching. This extends the Merton (1974) model to allow for regime switching risk.

Our second model captures the feature that default may occur anytime prior to or at maturity. This is an extension of the Black and Cox (1976) model to incorporate regime switching, and is more realistic.

We obtain **closed-form** analytic solutions for the value of the firm's equity and defaultable debt for both models.

## Appendix

*Proof of Lemma 5.2.1.* Consider the  $\mathbb{R}^N$ -valued process

$$Z_t := \exp\left(i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du\right) X_t.$$

Then

$$\begin{aligned} dZ_t &= \exp\left(i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du\right) dX_t + \exp\left(i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du\right) X_t \cdot i \langle \boldsymbol{\theta}, X_t \rangle X_t dt \\ &= (A + i \operatorname{diag} \boldsymbol{\theta}) Z_t dt + \exp\left(i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du\right) dM_t \end{aligned}$$

where the dynamics of  $X$  are given by (6.1). Thus

$$Z_t = X_0 + \int_0^t (A + i \operatorname{diag} \boldsymbol{\theta}) Z_u du + \int_0^t \exp\left(i \int_0^u \langle \boldsymbol{\theta}, X_s \rangle ds\right) dM_u.$$

The second term is a martingale, and so taking the expectation gives

$$\mathbb{E}[Z_t] = X_0 + \int_0^t (A + i \operatorname{diag} \boldsymbol{\theta}) \mathbb{E}[Z_u] du.$$

Solving this integral equation (which can be turned to an ODE), we have

$$\mathbb{E}[Z_t] = X_0 \exp((A + i \operatorname{diag} \boldsymbol{\theta})t).$$

Note that  $\mathbb{E}[Z_t]$  is an  $\mathbb{R}^N$ -valued process and the Fourier transform is, by definition, the characteristic function

$$\mathbb{E} \left[ \exp \left( i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du \right) \right].$$

Thus, the Fourier transform can be obtained by summing the components of  $\mathbb{E}[Z_t]$ . That is,

$$\begin{aligned} \mathbb{E} \left[ \exp \left( i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du \right) \right] &= \mathbb{E} \left[ \left\langle \exp \left( i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du \right), \mathbf{1} \right\rangle \right] \\ &= \langle X_0 \exp((A + i \operatorname{diag} \boldsymbol{\theta})t), \mathbf{1} \rangle. \end{aligned}$$

□

*Proof of Theorem 5.2.4.* We first note that, in general,

$$\int_0^t \langle \boldsymbol{\theta}, X_u \rangle du = \theta_1 \int_0^t \langle e_1, X_u \rangle du + \theta_2 \int_0^t \langle e_2, X_u \rangle du + \cdots + \theta_{N-1} \int_0^t \langle e_{N-1}, X_u \rangle du.$$

In particular, when  $N = 2$ , we have,

$$\int_0^t \langle \boldsymbol{\theta}, X_u \rangle du = \theta_1 \int_0^t \langle e_1, X_u \rangle du.$$

Note that  $X \in \{e_1, e_2\}$ , where  $e_1 = (0, 1)'$ , and  $e_2 = (1, 0)'$ . Consider the components of  $Z_t$ , denoted  $Z_t^1$  and  $Z_t^2$ . We have

$$\begin{aligned} Z_t^1 &= \left\langle \exp \left( i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du \right) X_t, e_1 \right\rangle \\ &= \exp \left( i \int_0^t \theta_1 \langle e_1, X_u \rangle du \right) \langle X_t, e_1 \rangle, \end{aligned}$$

and

$$\begin{aligned} Z_t^2 &= \left\langle \exp \left( i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du \right) X_t, e_2 \right\rangle \\ &= \exp \left( i \int_0^t \theta_1 \langle e_2, X_u \rangle du \right) \langle X_t, e_2 \rangle. \end{aligned}$$

Then, using Itô's lemma,

$$\begin{aligned}
Z_t^1 &= \langle e_1, X_0 \rangle + \int_0^t (i\theta_1) Z_u^1 du + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_1, AX_u du + dM_u \rangle \\
&= \langle e_1, X_0 \rangle + \int_0^t (i\theta_1) Z_u^1 du + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) (a_{11} \langle e_1, X_t \rangle - a_{22} \langle e_2, X_t \rangle) du \\
&\quad + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_1, dM_u \rangle \\
&= \langle e_1, X_0 \rangle + \int_0^t (i\theta_1 + a_{11}) Z_u^1 du - a_{22} \int_0^t Z_u^2 du + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_1, dM_u \rangle,
\end{aligned}$$

and

$$\begin{aligned}
Z_t^2 &= \langle e_2, X_0 \rangle + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_1, X_t \rangle \langle e_2, X_t \rangle + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_2, AX_u du + dM_u \rangle \\
&= \langle e_2, X_0 \rangle \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) (-a_{11} \langle e_1, X_t \rangle + a_{22} \langle e_2, X_t \rangle) du \\
&\quad + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_2, dM_u \rangle \\
&= \langle e_2, X_0 \rangle - a_{11} \int_0^t Z_u^1 du + a_{22} \int_0^t Z_u^2 du + \int_0^t \exp\left(i \int_0^u \theta_1 \langle e_1, X_s \rangle\right) \langle e_2, dM_u \rangle.
\end{aligned}$$

Taking expectations, we have

$$\mathbb{E}[Z_t^1] = \langle e_1, X_0 \rangle + \int_0^t (i\theta_1 + a_{11}) \mathbb{E}[Z_u^1] du - a_{22} \int_0^t \mathbb{E}[Z_u^2] du, \quad (5.5)$$

$$\mathbb{E}[Z_t^2] = \langle e_2, X_0 \rangle - a_{11} \int_0^t \mathbb{E}[Z_u^1] du + a_{22} \int_0^t \mathbb{E}[Z_u^2] du. \quad (5.6)$$

Therefore, solving the equations, we have

$$\mathbb{E}[Z_t^2] = e^{a_{22}t} \left( \langle e_2, X_0 \rangle - a_{11} \int_0^t e^{-a_{22}u} \mathbb{E}[Z_u^1] du \right).$$

Substituting into (5.5),

$$\mathbb{E}[Z_t^1] = \langle e_1, X_0 \rangle + \int_0^t (i\theta_1 + a_{11}) \mathbb{E}[Z_u^1] du - a_{22} \langle e_2, X_0 \rangle \int_0^t e^{a_{22}u} du + a_{22} a_{11} \int_0^t e^{a_{22}u} \left( \int_0^u e^{-a_{22}s} \mathbb{E}[Z_s^1] ds \right) du.$$

Writing  $Y_t^1 := e^{-a_{22}t} \mathbb{E}[Z_t^1]$ , then

$$\frac{dY_t^1}{dt} = (a_{11} - a_{22} + i\theta_1) Y_t^1 - a_{22} \langle e_2, X_0 \rangle + a_{22} a_{11} \int_0^t Y_u^1 du.$$

Therefore

$$\frac{dY_t^2}{dt} - (a_{11} - a_{22} + i\theta_1)\frac{dY_t^1}{dt} - a_{22}a_{11}Y_t^1 = 0$$

with

$$Y_0^1 = \mathbb{E}[Z_0^1] = \langle e_1, X_0 \rangle,$$

$$\left. \frac{dY_t^1}{dt} \right|_{t=0} = (a_{11} - a_{22} + i\theta_1)\langle e_1, X_0 \rangle - a_{22}\langle e_2, X_0 \rangle.$$

Suppose  $y_1$  and  $y_2$  are the roots of the quadratic equation

$$y^2 - (a_{11} - a_{22} + i\theta_1)y - a_{22}a_{11}y = 0. \quad (5.7)$$

Then

$$Y_t^1 = c_1 e^{y_1 t} + c_2 e^{y_2 t}$$

where  $c_1$  and  $c_2$  are determined by the above boundary conditions, so that

$$c_1 + c_2 = \langle e_1, X_0 \rangle, \quad (5.8)$$

$$c_1 y_1 + c_2 y_2 = (a_{11} - a_{22} + i\theta_1)\langle e_1, X_0 \rangle - a_{22}\langle e_2, X_0 \rangle. \quad (5.9)$$

Consequently

$$\mathbb{E}[Z_t^1] = e^{a_{22}t} Y_t^1 = c_1 e^{(a_{22}+y_1)t} + c_2 e^{(a_{22}+y_2)t}, \quad (5.10)$$

$$\begin{aligned} \mathbb{E}[Z_t^2] &= e^{a_{22}t} \left( \langle e_2, X_0 \rangle - a_{11} \int_0^t (c_1 e^{y_1 u} + c_2 e^{y_2 u}) du \right) \\ &= e^{a_{22}t} \langle e_2, X_0 \rangle - a_{11} e^{a_{22}t} \left( \frac{c_1}{y_1} (e^{y_1 t} - 1) + \frac{c_2}{y_2} (e^{y_2 t} - 1) \right). \end{aligned} \quad (5.11)$$

Thus the Fourier transform (characteristic function) is

$$\begin{aligned} E \left[ \exp \left( i \int_0^t \langle \boldsymbol{\theta}, X_u \rangle du \right) \right] &= \mathbb{E}[Z_t^1] + \mathbb{E}[Z_t^2] \\ &= \left( e^{a_{22}t} (c_1 e^{y_1 t} + c_2 e^{y_2 t}) + e^{a_{22}t} \langle e_2, X_0 \rangle - a_{11} e^{a_{22}t} \left( \frac{c_1}{y_1} (e^{y_1 t} - 1) + \frac{c_2}{y_2} (e^{y_2 t} - 1) \right) \right) \\ &= e^{a_{22}t} \left( c_1 e^{y_1 t} + c_2 e^{y_2 t} + \langle e_2, X_0 \rangle - a_{11} \left( \frac{c_1}{y_1} (e^{y_1 t} - 1) + \frac{c_2}{y_2} (e^{y_2 t} - 1) \right) \right) \end{aligned}$$

$$=e^{a_{22}t} \left( c_1 e^{y_1 t} \left( 1 - \frac{a_{11}}{y_1} \right) + c_2 e^{y_2 t} \left( 1 - \frac{a_{11}}{y_2} \right) + \langle e_2, X_0 \rangle + a_{11} \left( \frac{c_1}{y_1} + \frac{c_2}{y_2} \right) \right).$$

□

*Proof of Theorem 5.3.4.* Note that

$$A = \begin{bmatrix} -\lambda_1 & \lambda_2 \\ \lambda_1 & -\lambda_2 \end{bmatrix}.$$

Then the coupled homogeneous ODEs reduce to the following system

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2 P_1''(x) + \mu_1 x P_1'(x) - r_1 P_1(x) + c + \lambda_1 (P_2(x) - P_1(x)) = 0, \\ P_1(v_1) = \min(v_1, c/r), \\ \lim_{x \rightarrow \infty} P_1(x) = c/r. \end{array} \right.$$

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_2^2 P_2''(x) + \mu_2 x P_2'(x) - r_2 P_2(x) + c + \lambda_2 (P_1(x) - P_2(x)) = 0, \\ P_2(v_2) = \min(v_2, c/r), \\ \lim_{x \rightarrow \infty} P_2(x) = c/r. \end{array} \right.$$

From a practical perspective, we further assume that  $v_1 \leq v_2 \leq c/r$ , as the payoff to the bondholders at the default time never exceeds the value of the debt, discounted at a risk-free rate. Note that  $c/r$  is the value of the discounted default-free bond. Therefore

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_1^2 P_1''(x) + \mu_1 x P_1'(x) - r_1 P_1(x) + c + \lambda_1 (P_2(x) - P_1(x)) = 0 \\ P_1(v_1) = v_1 \\ \lim_{x \rightarrow \infty} P_1(x) = c/r. \end{array} \right. \quad (5.12)$$

$$\left\{ \begin{array}{l} \frac{1}{2}x^2\sigma_2^2 P_2''(x) + \mu_2 x P_2'(x) - r_2 P_2(x) + c + \lambda_2 (P_1(x) - P_2(x)) = 0 \\ P_2(v_2) = v_2 \\ \lim_{x \rightarrow \infty} P_2(x) = c/r. \end{array} \right. \quad (5.13)$$

Therefore, when  $v_1 \leq v_2 \leq c/r$ , we can write the coupled ODEs as follows. For  $x \in [0, v_1]$ , we have

$$P_1(x) = P_2(x) = x. \quad (5.14)$$

For  $x \in [v_1, v_2]$ , we have

$$\begin{cases} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(P_2(x) - P_1(x)) = 0 \\ P_2(x) = x. \end{cases} \quad (5.15)$$

For  $x \in [v_2, \infty)$ , we have

$$\begin{cases} \frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(P_2(x) - P_1(x)) = 0, \\ \frac{1}{2}x^2\sigma_2^2P_2''(x) + \mu_2xP_2'(x) - r_2P_2(x) + c + \lambda_2(P_1(x) - P_2(x)) = 0. \end{cases} \quad (5.16)$$

Now (5.16) has an characteristic function

$$g_1(\beta)g_2(\beta) = \lambda_1\lambda_2, \quad (5.17)$$

where

$$\begin{aligned} g_1(\beta) &= \lambda_1 + r_1 - \mu_1\beta - \frac{1}{2}\sigma_1^2\beta(\beta - 1), \\ g_2(\beta) &= \lambda_2 + r_2 - \mu_2\beta - \frac{1}{2}\sigma_2^2\beta(\beta - 1). \end{aligned}$$

This characteristic function has four distinct roots  $\beta_1 < \beta_2 < 0 < \beta_3 < \beta_4$ . To obtain a particular solution of (5.16), we consider

$$\begin{cases} -r_1P_1^* + c + \lambda_1(P_2^* - P_1^*) = 0 \\ -r_2P_2^* + c + \lambda_2(P_1^* - P_2^*) = 0. \end{cases}$$

This system reduces to

$$\begin{bmatrix} -(r_1 + \lambda_1) & \lambda_1 \\ -(r_2 + \lambda_2) & \lambda_2 \end{bmatrix} \begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix} = \begin{bmatrix} -c \\ -c \end{bmatrix},$$

and hence

$$\begin{aligned}
\begin{bmatrix} P_1^* \\ P_2^* \end{bmatrix} &= \begin{bmatrix} -(r_1 + \lambda_1) & \lambda_1 \\ -(r_2 + \lambda_2) & \lambda_2 \end{bmatrix}^{-1} \begin{bmatrix} -c \\ -c \end{bmatrix} \\
&= \frac{1}{r_2\lambda_1 - r_1\lambda_2} \begin{bmatrix} \lambda_2 & -\lambda_1 \\ r_2 + \lambda_2 & -(r_1 + \lambda_1) \end{bmatrix} \begin{bmatrix} -c \\ -c \end{bmatrix} \\
&= \frac{1}{r_2\lambda_1 - r_1\lambda_2} \begin{bmatrix} c(\lambda_1 - \lambda_2) \\ c((r_1 + \lambda_1) - (r_2 + \lambda_2)) \end{bmatrix},
\end{aligned}$$

Thus a particular solution of (5.16) is

$$\begin{aligned}
P_1^* &= \frac{c(\lambda_1 - \lambda_2)}{r_2\lambda_1 - r_1\lambda_2} \\
P_2^* &= \frac{c((r_1 + \lambda_1) - (r_2 + \lambda_2))}{r_2\lambda_1 - r_1\lambda_2}.
\end{aligned} \tag{5.18}$$

Therefore the general form of the solution to (5.16) is

$$\begin{aligned}
P_1(x) &= P_1^* + A_1x^{\beta_1} + A_2x^{\beta_2} + A_3x^{\beta_3} + A_4x^{\beta_4}, \\
P_2(x) &= P_2^* + B_1x^{\beta_1} + B_2x^{\beta_2} + B_3x^{\beta_3} + B_4x^{\beta_4},
\end{aligned}$$

with

$$B_i = l_i A_i, \tag{5.19}$$

and

$$l_i = l(\beta_i) = g_1(\beta_i)/\lambda_1 = \lambda_2/g_2(\beta_i) \tag{5.20}$$

When  $x \rightarrow \infty$ ,  $P_1(x)$  and  $P_2(x)$  are both bounded. Thus  $A_3 = A_4 = B_3 = B_4 = 0$ , and the solution is

$$\begin{aligned}
P_1(x) &= P_1^* + A_1x^{\beta_1} + A_2x^{\beta_2}, \\
P_2(x) &= P_2^* + B_1x^{\beta_1} + B_2x^{\beta_2}.
\end{aligned} \tag{5.21}$$

Next we solve (5.15). The first equation is

$$\frac{1}{2}x^2\sigma_1^2P_1''(x) + \mu_1xP_1'(x) - r_1P_1(x) + c + \lambda_1(x - P_1(x)) = 0. \quad (5.22)$$

This is an inhomogeneous equation, and thus the solution can be written as

$$P_1(x) = C_1x^{\gamma_1} + C_2x^{\gamma_2} + \phi(x), \quad (5.23)$$

where  $\phi(x)$  is a particular solution and  $\gamma_1$  and  $\gamma_2$  are the two roots of

$$\frac{1}{2}\sigma_1^2\gamma(\gamma - 1) + \mu_1\gamma - r_1 - \lambda_1 = 0. \quad (5.24)$$

To obtain  $\phi(x)$ , we assume that  $\phi(x) = ax + b$  and substitute it to equation (5.22). This yields

$$a\mu_1x - r_1(ax + b) + c - \lambda_1(ax + b) + \lambda_1x = 0,$$

or

$$x((\mu_1 - r_1 - \lambda_1)a + \lambda_1) + c - r_1b - \lambda_1b = 0.$$

Thus  $a = \lambda_1/(r_1 + \lambda_1 - \mu_1)$ ,  $b = c/(r_1 + \lambda_1)$ , and thus

$$\phi(x) = \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1}x + \frac{c}{r_1 + \lambda_1}. \quad (5.25)$$

Now we solve for the coefficients  $A_i$ ,  $B_i$ ,  $C_i$ . Using the boundary condition for  $P_2(x)$  (see(5.13), (5.16)) at  $v_2$  with  $x \in [v_2, \infty)$ ,

$$P_2^* + B_1v_2^{\beta_1} + B_2v_2^{\beta_2} = v_2,$$

or

$$l_1A_1v_2^{\beta_1} + l_2A_2v_2^{\beta_2} = v_2 - P_2^*, \quad (5.26)$$

Similarly, using the boundary and smoothness conditions for  $P_1(x)$  (see (5.12), (5.15)) at  $v_2$  with  $x \in [v_1, v_2]$ ,

$$\begin{cases} P_1^* + A_1v_2^{\beta_1} + A_2v_2^{\beta_2} = C_1v_2^{\gamma_1} + C_2v_2^{\gamma_2} + \phi(v_2), \\ \beta_1A_1v_2^{\beta_1} + \beta_2A_2v_2^{\beta_2} = \gamma_1C_1v_2^{\gamma_1} + \gamma_2C_2v_2^{\gamma_2} + v_2\phi'(v_2). \end{cases} \quad (5.27)$$



Using the boundary condition for  $P_1(x)$  (see (5.12), (5.15)) at  $v_1$  with  $x \in [v_1, v_2]$

$$C_1 v_1^{\gamma_1} + C_2 v_1^{\gamma_2} + \phi(v_1) = v_1, \quad (5.28)$$

From (5.27), we have

$$\begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix} \begin{bmatrix} C_1 v_2^{\gamma_1} \\ C_2 v_2^{\gamma_2} \end{bmatrix} = \begin{bmatrix} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \end{bmatrix},$$

or

$$\begin{bmatrix} C_1 v_2^{\gamma_1} \\ C_2 v_2^{\gamma_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \end{bmatrix}.$$

Therefore

$$\begin{aligned} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} &= \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \end{bmatrix} \\ &= \frac{1}{\gamma_2 - \gamma_1} \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} \begin{bmatrix} \gamma_2 & -1 \\ -\gamma_1 & 1 \end{bmatrix} \begin{bmatrix} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \end{bmatrix}. \end{aligned} \quad (5.29)$$

Using (5.29), we have

$$\begin{aligned} \begin{bmatrix} C_1 v_1^{\gamma_1} \\ C_2 v_1^{\gamma_2} \end{bmatrix} &= \begin{bmatrix} v_1^{\gamma_1} & 0 \\ 0 & v_1^{\gamma_1} \end{bmatrix} \begin{bmatrix} v_2^{-\gamma_1} & 0 \\ 0 & v_2^{-\gamma_2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \end{bmatrix} \\ &= \begin{bmatrix} v_1^{\gamma_1} v_2^{-\gamma_1} & 0 \\ 0 & v_1^{\gamma_1} v_2^{-\gamma_2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \\ \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_1} \\ \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \end{bmatrix}. \end{aligned} \quad (5.30)$$

Substituting (5.30) into (5.28), we have

$$\begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_1} \\ \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \end{bmatrix} = v_1 - \phi(v_1). \quad (5.31)$$

Writing out explicitly, we have

$$\begin{aligned} v_1 - \phi(v_1) &= \begin{bmatrix} 1 & 1 \end{bmatrix} \frac{1}{\gamma_2 - \gamma_1} \begin{bmatrix} \gamma_2 & -1 \\ -\gamma_1 & 1 \end{bmatrix} \begin{bmatrix} \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_1} \\ \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \end{bmatrix} \\ &= \frac{1}{\gamma_2 - \gamma_1} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_2 \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_1} - \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \\ -\gamma_1 \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} + \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \end{bmatrix}. \end{aligned}$$

Then

$$\begin{aligned} (\gamma_2 - \gamma_1) (v_1 - \phi(v_1)) &= \gamma_2 \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_1} - \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \\ &\quad - \gamma_1 \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} + \left( \beta_1 A_1 v_2^{\beta_1} + \beta_2 A_2 v_2^{\beta_2} - v_2 \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_2} \\ &= (\gamma_2 - \gamma_1) \left( P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2) \right) v_1^{\gamma_1} v_2^{-\gamma_1}. \end{aligned}$$

Thus

$$v_1^{-\gamma_1} v_2^{\gamma_1} (v_1 - \phi(v_1)) = P_1^* + A_1 v_2^{\beta_1} + A_2 v_2^{\beta_2} - \phi(v_2). \quad (5.32)$$

Combining (5.26) and (5.32), we obtain the system for  $A_1$  and  $A_2$

$$\begin{bmatrix} v_2^{\beta_1} & v_2^{\beta_2} \\ l_1 v_2^{\beta_1} & l_2 v_2^{\beta_2} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} v_1^{-\gamma_1} v_2^{\gamma_1} (v_1 - \phi(v_1)) - P_1^* + \phi(v_2) \\ v_2 - P_2^* \end{bmatrix}. \quad (5.33)$$

Solving the system (5.33), we have the solution for  $A_1$  and  $A_2$

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} v_2^{\beta_1} & v_2^{\beta_2} \\ l_1 v_2^{\beta_1} & l_2 v_2^{\beta_2} \end{bmatrix}^{-1} \begin{bmatrix} v_1^{-\gamma_1} v_2^{\gamma_1} (v_1 - \phi(v_1)) - P_1^* + \phi(v_2) \\ v_2 - P_2^* \end{bmatrix}. \quad (5.34)$$

Then, using  $B_i = l_i A_i$ , we can obtain coefficients  $B_1$  and  $B_2$ . Using (5.29), we can obtain coefficients  $C_1$  and  $C_2$ .

With these coefficients, the value functions  $P_1(x)$  and  $P_2(x)$  become

$$P_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} + \frac{c(\lambda_1 - \lambda_2)}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \phi(x) & \text{if } v_1 \leq x \leq v_2 \\ x & \text{if } x \leq v_1, \end{cases}$$

and

$$P_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} + \frac{c(r_1 + \lambda_1 - (r_2 + \lambda_2))}{r_2 \lambda_1 - r_1 \lambda_2} & \text{if } x > v_2, \\ x & \text{if } x \leq v_2, \end{cases}$$

where

$$\phi(x) = \frac{\lambda_1}{r_1 + \lambda_1 - \mu_1} x + \frac{c}{r_1 + \lambda_1}.$$

□

## Chapter 6

# Asset Price and Option Pricing with Markov Modulated Jump-Diffusion Dynamics

### 6.1 Introduction

Security markets modeled in continuous time have been a fundamental tool in modern financial theory. A basic example is the Black-Scholes-Merton model, where the dynamics of asset prices follow Geometric Brownian Motions; see Merton (1971, 1973, 1974), Black and Scholes (1973), Black and Cox (1976), et al. While this specification has achieved success in various areas of financial theory, using geometric Brownian motion to model asset prices entails several counterfactual empirical implications.

On the one hand, it has been widely documented that stock returns exhibit jumps. The importance of jumps arises not only from the time-series studies of stock prices, but also from cross-sectional studies of stock options (Bakshi et al., 1997; Bates, 2000).

On the other hand, there has been much empirical evidence suggesting that the aggregate economy is characterized by periodic shifts between distinct business cycles (see, e.g., Hamilton (1989)). In addition to the empirical evidence, there are also economic reasons to believe that regime switching is important in the understanding of the dynamics of the assets' prices. Empirically, the expansion and contraction periods in business cycles have potentially sizable effects on macroeconomic fundamentals, such as inflationary expectations, monetary policy, and nominal interest rates. Also, these regime shifts change corporations' growth prospects, and affect corporate earnings, as well as dividend payout decisions. On economic grounds, such switching in regimes gives rise to the possibility of significant impacts on firms' profitability, riskiness of assets and investments, and the growth rate of the

assets. In particular, Giesecke et al. (2011) conducted empirical analyses on the corporate bond market over the past 150 years. Their empirical data suggested that there are three regimes, and they also obtained the transition probabilities.

In this paper, we consider an asset price model with the aforementioned two features: jumps and regime shifts. In our model, the expected rate of return, and the jump intensities are both allowed to vary with different regimes. Moreover, we assume that, as is the case in reality, the underlying regimes are not directly observable. Rather, the regimes are inferred from the market observations, the time series of asset prices in our model.

Using hidden Markov chain and filtering techniques, we are able to estimate the current state of the underlying process from the market observations, namely the asset prices. Using the EM algorithm, we then obtain the maximum likelihood estimate for the parameters in our model.

Furthermore, our framework allows for a joint estimation of the parameters using time series data of both the asset's spot price, and the option price. Similar and related frameworks for joint estimation of parameters are proposed in, e.g., Pan (2002), Dai, Singleton and Yang (2007).

## 6.2 Model Specification

Let  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  be a complete, filtered probability space, with the filtration  $\mathbb{F} = (\mathcal{F})_{t \geq 0}$  satisfying the usual conditions. We suppose that the filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$  is rich enough to model the randomness of the observations process and of the state processes. Thus, all stochastic processes considered are, by definition,  $\mathbb{F}$ -adapted. We shall use this fact in various places in this paper. The probability measure  $\mathbb{P}$  is assumed to be the historical, “real world” measure.

Let  $X = (X_t)_{t \geq 0}$  be a finite state, continuous-time Markov chain, which represents the state of the economy is described the Markov chain  $X = \{X_t, t \geq 0\}$ . For notational

convenience, the state space of  $X$  can be taken to be, without loss of generality, the set of unit vectors  $\{e_1, e_2, \dots, e_N\}$ , where  $e_i = (0, \dots, 0, 1, 0, \dots, 0)' \in \mathbb{R}^N$  with the “1” in the  $i$ th entry and where the superscript  $'$  represents the transpose of a row vector. This is called the *canonical representation* of the state space of the Markov chain  $X$ . Suppose the process  $X$  is homogeneous in time.

Let  $A$  denote the constant rate matrix of the Markov chain  $X$ . Then, using the canonical representation of the state space, the dynamics of the Markov chain  $X$  have the following *semi-martingale representation* (see Elliott et al. (1995))

$$dX_t = AX_t dt + dM_t, \quad (6.1)$$

where  $(M_t)_{t \geq 0}$  is an  $\mathbb{R}^N$ -valued martingale with respect to the natural filtration generated by  $X$ .

In our model, the states of the Markov Chain  $X$  represent different economic environments. For example, there could be just two states for  $X$  representing “good” and “bad” economic regimes. The switching of the states of the economy can be attributed to structural changes in macroeconomic conditions, changes in political regimes and business cycles, etc. The states of the chain can also be interpreted as different characteristics of the issuers of the defaultable securities.

Suppose  $W$  is a standard Brownian motion on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We consider a risky asset, whose price  $S$  is modeled by the *Markov-modulated jump-diffusion*

$$\begin{aligned} \frac{dS_t}{S_{t-}} &= \mu_t dt + \sigma dW_t + dQ_t \\ &= \mu(X_t) dt + \sigma dW_t + dQ_t. \end{aligned} \quad (6.2)$$

Here, the expected rate of return  $\mu_t$  depends on the state of  $X_t$ , and the volatility  $\sigma$  is a constant. The jump component  $Q$  is a *Markov-modulated* compound poisson Process. The jump intensity and the distribution of the random jump size of  $Q$  both depend on the state of  $X_t$ .

Then, there exist vectors  $\mu = (\mu_1, \dots, \mu_N)'$  and  $\lambda = (\lambda_1, \dots, \lambda_N)'$  such that  $\mu_t = \langle \mu, X_t \rangle$  and  $\lambda_t = \langle \lambda, X_t \rangle$ . Consequently, the dynamics in equation (6.5) can be written simply as

$$\frac{dS_t}{S_{t-}} = \langle \mu, X_t \rangle dt + \sigma dW_t + dQ_t.$$

Let  $N$  denote a Poisson process with jump intensity  $\lambda_t = \langle \lambda, X_t \rangle$ , and unit jump size. Then  $Q_t = \int_0^t J_u dN_u$ , where  $J_u$  is the random jump size at time  $u$ . Equivalently, we have  $dQ_t = J_t dN_t$ .

In our model, the process  $X$  cannot be observed directly. Rather, what an investor observes is the price of the risky asset,  $S_t$ . For the sake of convenience, we consider  $y_t := \log S_t$ . Clearly, observations on  $S_t$  gives values for  $y_t$ , and vice versa.

**Proposition 6.2.1** *Suppose  $S$  satisfies the dynamics  $dS_t/S_{t-} = \mu_t dt + \sigma dW_t + dQ_t$ , where  $Q_t = \int_0^t J_u dN_u$ . Then  $y_t := \log S_t$  follows the dynamics*

$$dy_t = \left( \mu_t - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t + dZ_t,$$

where  $Z_t = \int_0^t \log(J_u + 1) dN_u$  is a compound Poisson process, and  $N$  has jump intensity  $\lambda_t$ .

*Proof.* See Appendix. □

With Proposition 6.2.1, we suppose that the observation comprises of  $y_t = \log S_t$ . The objective is to find the current state of  $X$  from the time-series of observations. Define  $\alpha_t := \mu_t - \frac{1}{2} \sigma^2$ . Then the observation process is of the form

$$y_t = \int_0^t \alpha_u du + \sigma W_t + Z_t$$

### 6.3 Filtering of the Current State

Define the families of  $\sigma$ -algebras  $\mathcal{F}_t^X := \sigma\{X_s : s \leq t\}$ ,  $\mathcal{F}_t^W := \sigma\{W_s, : s \leq t\}$  and  $\mathcal{F}_t^Z := \sigma\{Z_s : s \leq t\}$ . Let  $\mathcal{F}_t := \mathcal{F}_t^X \vee \mathcal{F}_t^W \vee \mathcal{F}_t^Z$ , and  $\mathcal{F}_t^y := \mathcal{F}_t^W \vee \mathcal{F}_t^Z$ .

Since we observe  $y_t$ , so we also observe the jump process  $y_t^d = Z_t$ , and therefore, we also know the continuous observation process  $y_t^c = y_t - y_t^d$ .

Recall that  $Z$  has jump intensity  $\lambda_t$ , and the distribution of the jump size  $Z_t$  is  $f_k(y)$ , when  $X_{u-} = e_k$ .

Suppose that under a “reference probability”  $\tilde{\mathbb{P}}$ , we have

- i)  $X$  is a finite state Markov chain with state space  $\mathbb{R}^K$  and rate matrix  $A$ .
- ii)  $y_t^c = \sigma \tilde{W}_t$ , where  $\tilde{W}_t$  is a Brownian motion under  $\tilde{\mathbb{P}}$ ; equivalently,  $\tilde{W}_t = \theta_t + W_t$ , where  $\theta_t = \alpha_t / \sigma$ .
- iii)  $y_t^d = Z_t$  is a jump process with intensity 1 and jump density  $f(y)$ ,  $y \in (-\infty, \infty)$ .

Define

$$\begin{aligned}\lambda_u &:= \sum_{k=1}^K \langle X_{u-}, e_k \rangle \lambda_k, \\ h_k(y) &:= \frac{\lambda_k f_k(y)}{f(y)}, \\ h_u(y) &:= \sum_{k=1}^K \langle X_{u-}, e_k \rangle h_k(y) = \sum_{k=1}^K \langle X_{u-}, e_k \rangle \frac{\lambda_k f_k(y)}{f(y)}.\end{aligned}$$

Also, let  $\hat{y}_u$  denote the realized jump size at time  $u$ , along the realized sample path  $\omega \in \Omega$ , which is determined by the observation process  $(y_t)_{t \geq 0}$  in our model.<sup>1</sup> Write

$$\begin{aligned}H_k(u)(\omega) &:= h_k(\hat{y}_u(\omega)), \\ H_u(\omega) &:= \sum_{k=1}^K \langle X_{u-}, e_k \rangle H_k(u)(\omega) = \sum_{k=1}^K \langle X_{u-}, e_k \rangle h_k(\hat{y}_u(\omega)).\end{aligned}$$

Define

$$\Lambda_t^c := \exp \left( - \int_0^t (-\theta_s) d\tilde{W}_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right)$$

---

<sup>1</sup>Note that in the above definitions,  $y$  is a random variable in  $(-\infty, \infty)$  denoting the random jump size at time  $u$ , while  $H_k(u)(\omega) = h_k(\hat{y}_u(\omega))$  is a deterministic function of  $u$  only, where  $\hat{y}_u$  is the realized jump size at time  $u$  along the realized sample path  $\omega \in \Omega$ . Therefore  $H_u = \sum_{k=1}^K \langle X_{u-}, e_k \rangle H_k(u)$  depends on the Markov chain  $X$  and time  $u$  only, once the realized sample path  $\omega \in \Omega$  is given.



$$\begin{aligned}
&= \exp \left( \int_0^t \theta_s dW_s + \frac{1}{2} \int_0^t \theta_s^2 ds \right) \\
&= \exp \left( \int_0^t \frac{\alpha_s}{\sigma^2} dy_s^c - \frac{1}{2} \int_0^t \frac{\alpha_s^2}{\sigma^2} ds \right), \\
\Lambda_t^d &:= 1 + \int_0^t \Lambda_{u-} \int_{-\infty}^{\infty} (h_u(y) - 1)(\gamma(dy, du) - f(y)dydu) \\
&= 1 + \int_0^t \Lambda_{u-} (H_u - 1) dN_u - \int_0^t \Lambda_{u-} (\lambda_u - 1) du,
\end{aligned}$$

A solution of  $\Lambda_t^d$  is

$$\begin{aligned}
\Lambda_t^d &= \exp \left( - \int_0^t (\lambda_u - 1) du + \int_0^t \log(H_u) dN_u \right) \\
&= \exp \left( - \int_0^t (\lambda_u - 1) du \right) \prod_{0 \leq u \leq t} (H_u)^{\Delta N_u}
\end{aligned}$$

We define the probability measure  $\mathbb{P}$  by setting

$$\left. \frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}} \right|_{\mathcal{F}_t} = \Lambda_t.$$

Then it is routine to show that (see, Elliott et al. (1995))

i)  $\Lambda_t$  is the unique solution to the integral equation

$$\Lambda_t = 1 + \int_0^t \Lambda_u \frac{\alpha_u}{\sigma^2} dy_u^c + \int_0^t \Lambda_{u-} (H_u - 1) dN_u - \int_0^t \Lambda_{u-} (\lambda_u - 1) du$$

ii) Under probability measure  $\mathbb{P}$ ,  $(y_t^c - \int_0^t \alpha_u du)/\sigma$  is a standard Brownian motion  $W = \{W_t, t \geq 0\}$ , i.e., under  $\mathbb{P}$ ,

$$y_t^c = \int_0^t \alpha_u du + \sigma W_t.$$

iii) Under probability measure  $\mathbb{P}$ ,  $y^d = Z = \{Z_t, t \geq 0\}$  is a compound Poisson process with compensator  $\int_0^t \lambda_u du$  and jump density  $f_k(y)$  if  $X_{u-} = e_k$ . That is,  $y_t^d - \int_0^t \lambda_u du$  is a  $(\mathbb{P}, \mathcal{F}_t)$  martingale. Therefore, under probability measure  $\mathbb{P}$

$$y_t = y_t^c + y_t^d = \int_0^t \alpha_u du + \sigma W_t + Z_t.$$

Using a version of Bayes' Theorem, we have

$$E[X_t|\mathcal{F}_t^y] = \frac{\tilde{\mathbb{E}}[\Lambda_t X_t|\mathcal{F}_t]}{\tilde{\mathbb{E}}[\Lambda_t|\mathcal{F}_t]}$$

Define  $q_t = \tilde{\mathbb{E}}[\Lambda_t X_t|\mathcal{F}_t^y]$ . Then

$$E[X_t|\mathcal{F}_t^y] = \frac{q_t}{\langle q_t, \mathbf{1} \rangle}$$

**Theorem 6.3.1** *The unnormalized filter  $\sigma(X)$  satisfies the vector stochastic equation*

$$\begin{aligned} q_t = & q_0 + \Lambda_0 X_0 + \int_0^t A\sigma(X_u)du + \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K)\sigma(X_u)dy_u^c \\ & + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1)\sigma(X_u)dN_u - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1)\sigma(X_u)du. \end{aligned}$$

*Proof.* By Itô's product rule

$$\begin{aligned} \Lambda_t X_t = & \Lambda_0 X_0 + \int_0^t X_{u-} d\Lambda_u + \int_0^t \Lambda_{u-} dX_u + [\Lambda, M]_t \\ = & \Lambda_0 X_0 + \int_0^t X_{u-} \Lambda_u \frac{\alpha_u}{\sigma^2} dy_u^c + \int_0^t X_{u-} \Lambda_{u-} (H_u - 1) dN_u \\ & - \int_0^t X_{u-} \Lambda_{u-} (\lambda_u - 1) du + \int_0^t \Lambda_{u-} A X_u du + \int_0^t \Lambda_{u-} dM_u \end{aligned}$$

for  $t \geq 0$ . Conditioning each side of the equation above on  $\mathcal{F}_t^y$ , we obtain

$$\begin{aligned} \sigma(X_t) = & \sigma(X_0) + \int_0^t X_{u-} d\Lambda_u + \int_0^t \Lambda_{u-} dX_u + [\Lambda, M] \\ = & \Lambda_0 X_0 + \int_0^t \tilde{\mathbb{E}}[X_{u-} \Lambda_u \alpha_u | \mathcal{F}_t^y] \frac{1}{\sigma^2} dy_u^c + \int_0^t \tilde{\mathbb{E}}[X_{u-} \Lambda_{u-} (H_u - 1) | \mathcal{F}_t^y] dN_u \\ & - \int_0^t \tilde{\mathbb{E}}[X_{u-} \Lambda_{u-} (\lambda_u - 1) | \mathcal{F}_t^y] du + \int_0^t A\sigma(X_u) du \\ = & \Lambda_0 X_0 + \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K)\sigma(X_u) dy_u^c + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1)\sigma(X_u) dN_u \\ & - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1)\sigma(X_u) du + \int_0^t A\sigma(X_u) du. \end{aligned}$$

This proves the desired equation. □

## 6.4 Parameter Estimation

In this section, we consider parameter estimation. We shall use the EM algorithm to estimate the transition matrix and the jump intensity  $\lambda_i$ .

To estimate  $\sigma$ , we note that the quadratic variation of  $y_t^c$  is  $\sigma^2$ . Consequently, an discretized approximations of the quadratic variation of  $y_t^c$  gives an estimate  $\hat{\sigma}$ .

Now the parameters to be estimated for the model are

$$\theta = \{a_{ji}, \alpha_i, \lambda_i : 1 \leq i, j, \leq K\}.$$

Note that, since  $\sum_{j=1}^K a_{ji} = 0$ , there is no need to estimate the diagonal elements  $a_{ii}$ ,  $i = 1, \dots, K$ .

In our parameter estimation scheme, since the distribution of the jump size  $f_k(y)$  is supposed known, the observation of the jump processes  $N$  and  $Z$  provide the same amount of information. The estimates of  $\theta$  involves the filtered estimates of the state  $X$ , two quantities related to the state  $X$ , as well as two quantities related to both  $X$  and  $N$ . We list the quantities to be estimated.

- 1)  $X_t$ , the state of the Markov chain  $X$ . The unnormalized estimate  $q_t$  has been obtain.
- 2) The occupation time  $O_t^i$  of the Markov chain in the state  $e_i$  over the time interval  $[0, t]$ , i.e.,

$$O_t^i = \int_0^t \langle X_u, e_i \rangle du, \quad 1 \leq i \leq K.$$

- 3) The number of jumps of Markov chain from state  $e_i$  to  $e_j$  over the time interval  $[0, t]$ ,  $J_t^{ij}$

$$J_t^{ij} = \int_0^t \langle X_{u-}, e_i \rangle \langle e_j, dX_u \rangle, \quad 1 \leq i, j \leq K.$$

4) The level integral  $G_{1t}^i$  with respect to  $y_t^c$

$$\begin{aligned} G_{1t}^i &= \int_0^t \langle X_u, e_i \rangle dy_u^c \\ &= \int_0^t \alpha_i \langle X_u, e_i \rangle du + \int_0^t \sigma \langle X_u, e_i \rangle dW_u, \quad 1 \leq i \leq K. \end{aligned}$$

5) The level integral  $G_{2t}^i$  with respect to the Poisson jump process  $N$

$$G_{1t}^i = \int_0^t \langle X_u, e_i \rangle dN_u, \quad 1 \leq i \leq K.$$

These quantities have semi-martingale decompositions (see Elliott et al. (1995) for details). For example, for  $J_t^{ij}$ , we have

$$\begin{aligned} J_t^{ij} &= \int_0^t \langle X_{u-}, e_i \rangle \langle e_j, dX_u \rangle \\ &= \int_0^t \langle X_{u-}, e_i \rangle \langle e_j, AX_u du + dM_u \rangle \\ &= \int_0^t \langle X_{u-}, e_i \rangle \langle e_j, AX_u \rangle du + M_t^{ij}, \end{aligned}$$

where

$$M_t^{ij} = \int_0^t \langle X_{u-}, e_i \rangle \langle e_j, dM_u \rangle.$$

Note that the integrand  $\langle X_{u-}, e_i \rangle e_j$  is predictable, and so  $M_t^{ij}$  is a martingale. Since  $\langle X_{u-}, e_i \rangle \langle e_j, AX_u \rangle = \langle X_{u-}, e_i \rangle a_{ji}$ , we have

$$J_t^{ij} = \int_0^t \langle X_{u-}, e_i \rangle a_{ji} du + M_t^{ij}, \quad 1 \leq i, j \leq K.$$

With these quantities, we shall use the EM algorithm to estimate  $\theta$ . Suppose  $\{\mathbb{P}_\theta, \theta \in \Theta\}$  is a family of probability measures on  $\{\mathbb{P}, \mathcal{F}\}$ , all absolutely continuous with respect to a fixed probability measure  $\mathbb{P}$ . The likelihood function of the parameter  $\theta$ , based on  $\mathcal{F}_t^y$ , is

$$\mathcal{L}(\theta) = \mathbb{E} \left[ \frac{d\mathbb{P}_\theta}{d\mathbb{P}} \middle| \mathcal{F}_t^y \right].$$

Suppose the model is first determined by a set of parameters  $\theta = \{a_{ji}, \alpha_i, \lambda_i : 1 \leq i, j, \leq K\}$ , and we wish to determine a new set of parameters  $\hat{\theta} = \{\hat{a}_{ji}, \hat{\alpha}_i, \hat{\lambda}_i : 1 \leq i, j, \leq K\}$ , which maximizes the conditional pseudo log-likelihood defined below. To change the parameters  $\theta$  to  $\hat{\theta}$ , we define

$$\begin{aligned} \frac{d\mathbb{P}_{\theta'}}{d\mathbb{P}_{\theta}} \Big| \mathcal{F}_t := & \prod_{i \neq j} L_t^{ij} \exp \left( \int_0^t \frac{1}{\sigma^2} (\alpha'_u - \alpha_u) du - \frac{1}{2} \int_0^t \frac{1}{\sigma^2} (\alpha_u'^2 - \alpha_u^2) du \right. \\ & \left. - \int_0^t (\lambda'_u - \lambda_u) du + \int_0^t (\log(\lambda'_u) - \log(\lambda_u)) dN_u \right) \end{aligned}$$

with

$$L_t^{ij} = \left( \frac{a'_{ji}}{a_{ji}} \right)^{J_t^{ij}} \exp \left( \int_0^t (a'_{ji} - a_{ji}) \langle X_u, e_i \rangle du \right),$$

and compute the quantity

$$\mathbb{E} \left[ \log \frac{d\mathbb{P}_{\theta'}}{d\mathbb{P}_{\theta}} \Big| \mathcal{F}_t \right].$$

Taking the log, we have

$$\begin{aligned} \log \left( \frac{d\mathbb{P}_{\theta'}}{d\mathbb{P}_{\theta}} \Big| \mathcal{F}_t \right) &= \sum_{i \neq j} \log(L_t^{ij}) + \int_0^t \frac{1}{\sigma^2} (\alpha'_u - \alpha_u) du - \frac{1}{2} \int_0^t \frac{1}{\sigma^2} (\alpha_u'^2 - \alpha_u^2) du \\ &\quad - \int_0^t (\lambda'_u - \lambda_u) du + \int_0^t (\log(\lambda'_u) - \log(\lambda_u)) dN_u \\ &= \sum_{i \neq j} \left( J_t^{ij} (a'_{ji} - a_{ji}) + \int_0^t (a'_{ji} - a_{ji}) \langle X_u, e_i \rangle du \right) \\ &\quad + \int_0^t \frac{1}{\sigma^2} (\alpha'_u - \alpha_u) du - \frac{1}{2} \int_0^t \frac{1}{\sigma^2} (\alpha_u'^2 - \alpha_u^2) du \\ &\quad - \int_0^t (\lambda'_u - \lambda_u) du + \int_0^t (\log(\lambda'_u) - \log(\lambda_u)) dN_u. \end{aligned}$$

Therefore, taking the conditional expectation on  $\mathcal{F}_t^y$  of both sides, we obtain

$$\begin{aligned} \mathbb{E} \left[ \log \frac{d\mathbb{P}_{\theta'}}{d\mathbb{P}_{\theta}} \Big| \mathcal{F}_t \right] &= \sum_{i \neq j} \left( \hat{J}_t^{ij} \log a'_{ji} - a'_{ji} \hat{O}_t^i \right) + \sum_{i=1}^K \frac{1}{\sigma^2} \left( \alpha_i \hat{G}_{1t}^i - \frac{\alpha_i^2}{2} \hat{O}_t^i \right) \\ &\quad + \sum_{i=1}^K \left( \hat{G}_{2t}^i \log(\lambda_i) - \lambda_i \hat{O}_t^i \right) + R(\theta). \end{aligned}$$

Here the term  $R(\theta)$  does not involve the parameter  $\theta'$ , and so the E-step consists of computing  $\widehat{J}_t^{ij}$ ,  $\widehat{O}_t^i$ ,  $\widehat{G}_{1t}^i$ , and  $\widehat{G}_{2t}^i$ .

For the M-step, set the partial derivatives with respect to  $a'_{ji}$ ,  $\alpha'_i$ , and  $\lambda'_i$  to zero. This yields estimates of  $\widehat{\theta}$  as

$$\begin{aligned}\widehat{a}_{ji} &= \frac{\widehat{J}_t^{ij}}{\widehat{O}_t^i} = \frac{\sigma(J_t^{ij})}{\sigma(O_t^i)}, \\ \widehat{\alpha}_i &= \frac{\widehat{G}_{1t}^i}{\widehat{O}_t^i} = \frac{\sigma(G_{1t}^i)}{\sigma(O_t^i)}, \\ \widehat{\lambda}_i &= \frac{\widehat{G}_{2t}^i}{\widehat{O}_t^i} = \frac{\sigma(G_{2t}^i)}{\sigma(O_t^i)}.\end{aligned}$$

We now derive filters for  $J_t^{ij}$ ,  $O_t^i$ ,  $G_{1t}^i$ , and  $G_{2t}^i$ , which are needed to compute estimates  $\widehat{a}_{ji}$ ,  $\widehat{\alpha}_i$ , and  $\widehat{\lambda}_i$ . Usually the quantities  $\sigma(J_t^{ij})$ ,  $\sigma(O_t^i)$ ,  $\sigma(G_{1t}^i)$ , and  $\sigma(G_{2t}^i)$  can not be computed recursively. We consider instead the quantities  $\sigma(J_t^{ij} X_t)$ ,  $\sigma(O_t^i X_t)$ ,  $\sigma(G_{1t}^i X_t)$ , and  $\sigma(G_{2t}^i X_t)$ . Once these vector quantities are obtained,  $\sigma(J_t^{ij})$ ,  $\sigma(O_t^i)$ ,  $\sigma(G_{1t}^i)$ , and  $\sigma(G_{2t}^i)$  can be calculated through normalization. For example,

$$\sigma(J_t^{ij}) = \sum_{i=1}^K \langle \sigma(J_t^{ij} X_t), e_i \rangle = \langle \sigma(J_t^{ij} X_t), \mathbf{1} \rangle.$$

**Theorem 6.4.1** For  $t \geq 0$  and  $1 \leq i, j \leq K$ ,  $i \neq j$ ,

$$\begin{aligned}\sigma(J_t^{ij} X_t) &= \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K) \sigma(J_u^{ij} X_u) dy_u^c + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1) \sigma(J_u^{ij} X_u) dN_u \\ &\quad - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1) \sigma(J_u^{ij} X_u) du + \int_0^t \langle \sigma(X_u), e_i \rangle a_{ij} e_j du + \int_0^t A \sigma(J_u^{ij} X_u) du.\end{aligned}$$

*Proof.* Using Itô's product rule, as well as the semi-martingale representation of the Markov chain  $M$  and the semi-martingale decomposition of  $J_t^{ij}$ , we have

$$\begin{aligned}J_t^{ij} X_t &= \int_0^t X_u \langle X_{u-}, e_i \rangle a_{ij} du + \int_0^t X_{u-} dM_u^{ij} + \int_0^t A X_u J_u^{ij} du \\ &\quad + \int_0^t J_{u-}^{ij} dM_u + \sum_{0 \leq u \leq t} (e_j - e_i) \langle X_{u-}, e_i \rangle \langle X_u, e_j \rangle\end{aligned}$$

$$= \int_0^t \langle X_u, e_i \rangle a_{ij} e_j du + \int_0^t A J_u^{ij} X_u du + \int_0^t \zeta_{u-} dM_u,$$

for some predictable process  $\zeta$ . Then again use Itô's product rule, we have

$$\begin{aligned} \Lambda_t J_t^{ij} X_t &= \Lambda_0 J_0^{ij} X_0 + \int_0^t J_u^{ij} X_u d\Lambda_u + \int_0^t \Lambda_u \langle X_u, e_i \rangle a_{ij} e_j du \\ &\quad + \int_0^t \Lambda_u A J_u^{ij} X_u du + \int_0^t \Lambda_u \zeta_{u-} dM_u \\ &= \int_0^t J_u^{ij} X_u \Lambda_u \frac{\alpha_u}{\sigma^2} dy_u^c + \int_0^t J_u^{ij} X_u \Lambda_{u-} (H_u - 1) dN_u - \int_0^t J_u^{ij} X_u \Lambda_{u-} (\lambda_u - 1) du \\ &\quad + \int_0^t \Lambda_u \langle X_u, e_i \rangle a_{ij} e_j du + \int_0^t \Lambda_u A J_u^{ij} X_u du + \int_0^t \Lambda_u \zeta_{u-} dM_u. \end{aligned}$$

Conditioning each side on  $\mathcal{F}_t^y$  under the probability measure  $\tilde{\mathbb{P}}$ , we have

$$\begin{aligned} \sigma(J_t^{ij} X_t) &= \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K) \sigma(J_u^{ij} X_u) dy_u^c + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1) \sigma(J_u^{ij} X_u) dN_u \\ &\quad - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1) \sigma(J_u^{ij} X_u) du + \int_0^t \langle \sigma(X_u), e_i \rangle a_{ij} e_j du + \int_0^t A \sigma(J_u^{ij} X_u) du. \end{aligned}$$

The desired equation is proven.  $\square$

**Theorem 6.4.2** For  $t \geq 0$  and  $1 \leq i, j \leq K$ ,  $i \neq j$ ,

$$\begin{aligned} \sigma(O_t^i X_t) &= \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K) \sigma(O_u^i X_u) dy_u^c + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1) \sigma(O_u^i X_u) dN_u \\ &\quad - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1) \sigma(O_u^i X_u) du + \int_0^t \langle \sigma(X_u), e_i \rangle e_i du + \int_0^t A \sigma(O_u^i X_u) du, \\ \sigma(G_{1t}^i X_t) &= \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K) \sigma(G_{1u}^i X_u) dy_u^c + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1) \sigma(G_{1u}^i X_{u-}) dN_u \\ &\quad - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1) \sigma(G_{1u}^i X_{u-}) du + \int_0^t \langle \sigma(X_u), e_i \rangle \alpha_i e_i du + \int_0^t A \sigma(G_{1u}^i X_u) du, \\ \sigma(G_{2t}^i X_t) &= \int_0^t \frac{1}{\sigma^2} \text{diag}(\alpha_1, \dots, \alpha_K) \sigma(G_{2u}^i X_u) dy_u^c + \int_0^t \text{diag}(H_1(u) - 1, \dots, H_K(u) - 1) \sigma(G_{2u-}^i X_{u-}) dN_u \\ &\quad - \int_0^t \text{diag}(\lambda_1 - 1, \dots, \lambda_K - 1) \sigma(G_{2u-}^i X_{u-}) du + \int_0^t \langle \sigma(X_u), e_i \rangle \lambda_i e_i du + \int_0^t A \sigma(G_{2u}^i X_u) du. \end{aligned}$$

## 6.5 Applications on Option Pricing

In this section, we explore the application of our asset price model on option pricing. Recall that for  $y_t = \log S_t$ ,

$$dy_t = \left( \mu_t - \frac{1}{2}\sigma^2 \right) dt + \sigma dW_t + dZ_t,$$

where  $Z_t = \int_0^t \log(J_u + 1) dN_u$  is a compound Poisson process, and  $N$  has jump intensity  $\lambda_t$ . Equivalently,

$$\begin{aligned} y_t &= \int_0^t \left( \mu_u - \frac{1}{2}\sigma^2 \right) du + \int_0^t \sigma dW_u + Z_t \\ &= \int_0^t \left( \mu_u + \lambda_u \mathbb{E}[\delta_u] - \frac{1}{2}\sigma^2 \right) du + \int_0^t \sigma dW_u + \int_0^t (\delta_u dN_u - \lambda_u \mathbb{E}[\delta_u] du) \\ &= \int_0^t \left( \beta_u - \frac{1}{2}\sigma^2 \right) du + \int_0^t \sigma dW_u + \int_0^t (\delta_u dN_u - \lambda_u \mathbb{E}[\delta_u] du). \end{aligned}$$

Here the drift term and the jump intensity are both switching with the Markov chain  $X$ .

Having estimated the parameters from market observations, an immediate question is can we use the model to price certain derivatives on the asset. We now indicate how options on the asset can be priced, with the hidden Markov-modulated jump-diffusion dynamics for the asset.

In Elliott et al. (2007), an approach using the Esscher transform is introduced which we now follow. Let

$$\Lambda_t^1 := \exp \left( \int_0^t \theta_{1u} \sigma dW_u - \frac{1}{2} \int_0^t \theta_{1u}^2 \sigma^2 du \right). \quad (6.3)$$

Also, denote the distribution of jump size  $y$  by  $f(y)$ , and define

$$\Lambda_t^2 := \exp \left( \int_0^t \int_{-\infty}^{\infty} (1 - e^{-\theta_{2u}y}) \lambda_u f(y) dy du - \int_0^t \theta_{2u} \delta_u dN_u \right). \quad (6.4)$$

Using  $\Lambda_t := \Lambda_t^1 \Lambda_t^2$  as the Radon-Nikodym derivative  $d\mathbb{Q}/d\mathbb{P}$ , we can obtain infinitely many equivalent martingale measures.



### 6.5.1 Idiosyncratic Jump Risk

To obtain a reasonably closed-form expression, we first consider the case when the jump risk is idiosyncratic, so that it can be completely hedged away. This is similar to the case for the jump-diffusion asset dynamics considered in Merton (1974), where the diffusion and jump components are assumed to represent the systematic and idiosyncratic risk, respectively, under the assumption that the two components are independent. Therefore, as in Merton (1974), no risk premium is attached to jumps and hence, the jump risk is not priced. In mathematical terms, the risk neutral properties of the jump component of  $S_t$  are the same as its statistical (actuarial) properties. In particular, the distributions of jump times and jump sizes are unchanged between the physical (real-world) and the risk neutral measures.

In this case, since the jump risk does not need to be priced, we set  $\theta_2 = 0$ . With the assumption that  $\theta_2 = 0$ , one can then derive the coupled PIDE for the price of a European option  $V = (V_1, \dots, V_N)$ , where  $V_i := V(t, T, S, e_i)$ :

$$\begin{aligned} & -r_t V + \frac{\partial V}{\partial S} + \frac{\partial V}{\partial S} S_{t-} \left( r_t + \int_{-\infty}^{\infty} (e^y - 1 - y) \lambda_t f(y) dy \right) + \frac{1}{2} \sigma^2 S_{t-}^2 \frac{\partial^2 V}{\partial S^2} \\ & + \int_{-\infty}^{\infty} \left( V(t, T, S_{t-} + e^y - 1, X_{t-}) - V(t, T, S_{t-}, X_{t-}) - \frac{\partial V}{\partial S} (e^y - 1) \right) \lambda_t f(y) dy + \langle V, AX_{t-} \rangle = 0, \end{aligned}$$

with terminal condition

$$V(T, T, S, X) = V(S_T).$$

This is the coupled PIDE for the price of a European option in the hidden Markov-modulated jump diffusion model. Detailed calculation can be found in Elliott et al. (2007).

### 6.5.2 Systematic Jump Risk

When jumps represent a systematic risk, such risk is not diversifiable, and thus cannot be hedged away. If this is the case, the jump risk needs to be priced. This corresponds to the case that  $\theta_2 \neq 0$  in equation (6.4). Then the Radon-Nikodym derivative is  $d\mathbb{Q}/d\mathbb{P} = \Lambda_t^1 \Lambda_t^2$ , and the pricing kernel is  $\exp\left(-\int_0^t r_u du\right) \Lambda_t^1 \Lambda_t^2$ . There are infinitely many risk neutral measures

(equivalent martingale measures)  $\mathbb{Q}$ . For the one-regime jump-diffusion dynamics, Colwell and Elliott (1993) derived an integro differential equation. Colwell and Elliott (1993) also considered a quadratic hedging approach to determine the risk-neutral, pricing measure  $\mathbb{Q}$ .

## 6.6 Concluding Remarks.

In this paper, we consider a filtering problem of a **Markov-modulated** jump-diffusion model. We obtain the filter for the current state.

We also consider estimation of the parameters of the model. By using the EM algorithm, we obtain estimates which converge to the maximum likelihood estimates.

Applications of our asset dynamics model on option pricing are also discussed.

## Appendix

Let  $N$  denote the standard Poisson process with jump intensity  $\lambda_t = \langle \lambda, X_t \rangle$ , and unit jump size. Suppose the price of a risky asset at time  $t$  is  $S_t$ , which follows the dynamics

$$\begin{aligned} dS_t/S_{t-} &= \mu_t dt + \sigma dW_t + dQ_t \\ &= \langle \mu, X_t \rangle dt + \sigma dW_t + dQ_t. \end{aligned} \tag{6.5}$$

Here  $\mu_t = \langle \mu, X_t \rangle$ ,  $\sigma$  is a constant, and  $Q$  is a compound poisson process, with jump intensity  $\lambda_t = \langle \lambda, X_t \rangle$ . Then

$$Q_t = \int_0^t J_u dN_u.$$

where  $J_u$  is the jump size at time  $u$ .

The solution of  $S_t$  is given by

$$S_t = S_0 \exp \left( \int_0^t (\mu_u - \frac{1}{2} \sigma^2) du + \sigma W_t \right) \prod_{u=0}^t (J_u + 1)^{\Delta N_u}$$

The observation is  $y_t = \log S$ . Suppose  $\delta_u = \log(J_u + 1)$ . Then

$$y_t = \int_0^t (\mu_u - \frac{1}{2}\sigma^2)du + \int_0^t \delta_u dN_u$$

Define

$$Z_t = \int_0^t \delta_u dN_u = \int_0^t \log(J_u + 1)dN_u.$$

Then,  $Z$  is a compound Poisson process with jump intensity  $\lambda_t$  (the same as  $N$ ), and jump size  $\log(J_u + 1)$ , where  $J_u$  is the jump size of the compound Poisson process  $Q$ . Thus

$$\begin{aligned} y_t = \log S_t &= \int_0^t (\mu_u - \frac{1}{2}\sigma^2)du + \sigma W_t + \int_0^t \delta_u dN_u \\ &= \int_0^t (\mu_u - \frac{1}{2}\sigma^2)du + \sigma W_t + Z_t. \end{aligned}$$

## Appendix: List of Publications

The thesis is based on the author's joint work with Professor Robert Elliott. A list of the publications of the author is included below.

- [1 ] Robert J. Elliott and Jia Shen, General Equilibrium Pricing with Multiple Dividend Streams and Regime Switching, *Quantitative Finance*, **15** (9), 1543–1557, 2015.
- [2 ] Robert J. Elliott and Jia Shen, Credit Risk with Self-Exciting Processes, *Annals of Finance*, Forthcoming. (accepted Nov 23, 2014) DOI:10.1007/s10436-015-0259-z.
- [3 ] Robert J. Elliott and Jia Shen, Dynamic Optimal Capital Structure with Regime Switching Risk, *Annals of Finance*, **11** (2), 199–220, 2015.
- [4 ] Robert J. Elliott and Jia Shen, Asset Price Dynamics and Option Pricing with Markov Modulated Jump-Diffusion, submitted.
- [5 ] Robert J. Elliott and Jia Shen, Credit risk with latent contagion and frailty: default probabilities, pricing, and hedging, submitted.
- [6 ] Robert J. Elliott and Jia Shen, Credit Spread of Defaultable Corporate Bonds with Regime Switching, submitted.

## Bibliography

- [1] Mas-Colell A., M.D. Whinston, and J.R. Green. *Microeconomic Theory*. Oxford University Press, 1995.
- [2] S. Altug and P. Labadie. *Asset Pricing for Dynamic Economies*. Cambridge University Press, New York, 2008.
- [3] G. Amisano and O. Tristani. A DSGE model of the term structure with regime shifts. European central bank working paper series, 2010.
- [4] S. Azizpour, K. Giesecke, and G. Schwenkler. Exploring the sources of default clustering. working paper, Stanford University, 2014.
- [5] K. Back. *Asset Pricing and Portfolio Choice Theory*. Oxford University Press, New York, 2010.
- [6] G.S. Bakshi and Z. Chen. Inflation, asset prices, and the term structure of interest rates in monetary economics. *Review of Financial Studies*, 9:241–275, 1996.
- [7] T.R. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation and Hedging*. Springer-Verlag, 2002.
- [8] A. Buraschi, P. Porchia, and F. Trojani. The cross-section of expected stock returns: learning about distress and predictability in heterogeneous orchards. working paper, 2010.
- [9] Stephen G. Cecchetti, Pok sang Lam, and Nelson C. Mark. The equity premium and the risk-free rate: matching the moments. *Journal of Monetary Economics*, 31(1):21–45, 1993.
- [10] S. Chib, K.H. Kang, and S. Ramamurthy. Term structure of interest rates in a dsge model with regime changes. working paper, Washington University in St. Louis, 2010.

- [11] J.H. Cochrane, F.A. Longstaff, and P. Santa-Clara. Two trees. *The Review of Financial Studies*, 21(1):347–385, 2008.
- [12] P. Collin-Dufresne, R. Goldstein, and J. Helwege. Is credit event risk priced? modeling contagion via the updating of beliefs. preprint, Carnegie Mellon University, 2003.
- [13] J. C. Cox and C.-F. Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49:33–83, 1989.
- [14] J.C. Cox, J.E. Ingersoll Jr., and S.A. Ross. An analysis of variable rate loan contracts. *The Journal of Finance*, 35(2):389–403, 1980.
- [15] J.C. Cox, J.E. Ingersoll Jr., and S.A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384, 1985.
- [16] J.C. Cox, J.E. Ingersoll Jr., and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.
- [17] J. Cvitanić, J. Ma, and J. Zhang. Laws of large numbers for self-exciting correlated defaults. *Stochastic Processes and Applications*, 122:2781–2810, 2011.
- [18] Q. Dai, K.J. Singleton, and W. Yang. Regime shifts in a dynamic term structure model of u.s. treasury bond yields. *Review of Financial Studies*, 20:1669–1706, 2007.
- [19] M.H.A. Davis. Option pricing in incomplete markets. In M.A.H. Dempster and S.R. Pliska, editors, *Mathematics of Derivative Securities*. Cambridge University Press, 1998.
- [20] D. Duffie, A. Eckner, G. Horel, and L. Saita. Frailty correlated default. *Journal of Finance*, 64:2089–2123, 2009.
- [21] D Duffie and L.G. Epstein. Asset pricing with stochastic differential utility. *The Review of Financial Studies*, 5:411–436, 1992.

- [22] D. Duffie and D. Lando. Term structure of credit spreads with incomplete accounting information. *Econometrica*, 69:633–664, 2001.
- [23] D. Duffie, M. Schroder, and C. Skiadas. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *Annals of Applied Probability*, 6:1075–1090, 1996.
- [24] D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12:687–720, 1999.
- [25] D. Duffie and K. Singleton. *Credit Risk: Pricing, Measurement and Management*. Princeton University Press, New Jersey, 2003.
- [26] D. Duffie and C. Skiadas. Continuous-time security pricing: a utility gradient approach. *Journal of Mathematical Economics*, 23:107–131, 1994.
- [27] D. Duffie and W. Zame. The consumption-based capital asset pricing model. *Econometrica*, 57:1279–1297, 1989.
- [28] R.J. Elliott, L. Aggoun, and J.B. Moore. *Hidden Markov models: Estimation and control*. Applications of Mathematics (New York) 29. Springer-Verlag, New York, 1995.
- [29] R.J. Elliott and H. Föllmer. Orthogonal martingale representation. In *Stochastic Analysis*, pages 139–152. Academic Press, Boston, MA, 1991.
- [30] R.J. Elliott, M. Jeanblanc, and M. Yor. On models of default risk. *Mathematical Finance*, page 179195, 2000.
- [31] R.J. Elliott and R.S. Mamon. A complete yield curve description of a markov interest rate model. *Int. J. Theor. Appl. Finance*, 6:317–326, 2003.
- [32] H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. In M.H.A. Davis and R.J. Elliott, editors, *Applied Stochastic Analysis*, pages

- 389–414. Gordon and Breach, London, 1990.
- [33] H. Föllmer and D. Sondermann. Hedging of non-redundant contingent-claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics*, pages 205–223. North-Holland, Amsterdam, 1986.
- [34] R. Frey and W. Runggaldier. Pricing credit derivatives under incomplete information: A nonlinear filtering approach. *Finance and Stochastics*, 14:495–526, 2010.
- [35] R. Frey and T. Schmidt. Pricing and hedging of credit derivatives via the innovations approach to nonlinear filtering. *Finance and Stochastics*, 16:105–133, 2012.
- [36] K. Giesecke, F. Longstaff, S. Schaefer, and I. Strebulaev. Corporate bond default risk: a 150-year perspective. *Journal of Financial Economics*, 102:233–250, 2011.
- [37] K. Giesecke, F. Longstaff, S. Schaefer, and I. Strebulaev. Macroeconomic effects of corporate default crises: a long-term perspective. *Journal of Financial Economics*, 111:297–310, 2014.
- [38] G. Graziano and C. Rogers. A dynamic approach to the modelling of correlation credit derivatives using markov chains. *Int. J. Theor. Appl. Finance*, 12:45–62, 2009.
- [39] B. Hajek and E. Wong. *Stochastic Processes in Engineering Systems*. Springer-Verlag, New York, 1985.
- [40] J. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.
- [41] J. Heaton and D.J. Lucas. Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy*, 104(3):443–487, 1996.
- [42] Chi-Fu Huang. An intertemporal general equilibrium asset pricing model: the case of diffusion information. *Econometrica*, 55:117–142, 1987.



- [43] R.A. Jarrow, D. Lando, and S.M. Turnbull. A markov model for the term structure of credit risk spreads. *Review of financial studies*, 10(2):481–523, 1997.
- [44] R.A. Jarrow and S.M. Turnbull. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50:53–86, 1995.
- [45] I. Karatzas, J. Lehoczky, and S. E. Shreve. Optimal portfolio and consumption decisions for a ‘small investor’ on a finite horizon. *SIAM Journal of Control and Optimization*, 25:1557–1586, 1987.
- [46] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, Berlin, 1988.
- [47] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag, New York, 1992.
- [48] D. Lando. On cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120, 1998.
- [49] D. Lando. *Credit Risk Modeling: Theory and Applications*. Princeton University Press, New Jersey, 2004.
- [50] Y. Lengwiler. *Microfoundations of Financial Economics*. Princeton University Press, 2004.
- [51] S.F. and J. Werner LeRoy. *Principles of Financial Economics*. Cambridge University Press, 2001.
- [52] Deborah J. Lucas. Asset pricing with undiversifiable income risk and short sales constraints: Deepening the equity premium puzzle. *Journal of Monetary Economics*, 34:325–341, 1994.
- [53] R.E. Lucas. Asset prices in an exchange economy. *Econometrica*, 46:1429–1445, 1978.

- [54] I. Martin. The lucas orchard. *Econometrica*, 81(1):55–111, 2013.
- [55] R. Mehra and E.C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15:145–161, 1985.
- [56] R. Mehra and E.C. Prescott. *The Equity Premium in Retrospect*, volume 1B. Elsevier, North Holland, 2003.
- [57] C.A. Parlour, R. Stanton, and J. Walden. Revisiting asset pricing puzzles in an exchange economy. *The Review of Financial Studies*, 24(3):629–674, 2011.
- [58] G. Pennacchi. *Theory of Asset Pricing*. Prentice Hall, 2007.
- [59] S. Pliska. A stochastic calculus model of continuous trading: optimal portfolios. *Mathematics of Operations Research*, 11:371–382, 1986.
- [60] P. Schönbucher. *Credit derivatives pricing models: models, pricing and implementation*. Wiley, Chichester, UK, 2003.
- [61] P. Schönbucher. Information driven default contagion. Working paper, ETH, Zurich, 2003.
- [62] C. Skiadas. *Asset Pricing Theory*. Princeton University Press, 2009.