Two Essays in Financial Econometrics

A Dissertation Presented

by

Yang YU

to

The Graduate School in Partial Fulfillment of the Requirements for the

Degree of

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

December 2013

Stony Brook University

The Graduate School

Yang YU

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Haipeng Xing - Dissertation Advisor
Associate Professor, Applied Mathematics and Statistics, SUNYSB

Wei Zhu - Chairperson of Defense
Professor & Deputy Chair, Applied Mathematics and Statistics, SUNYSB

Svetlozar Rachev
Professor of Finance, College of Business, SUNYSB

Yixin Fang
Assistant Professor, Division of Biostatistics, School of Medicine,
New York University

This dissertation is accepted by the Graduate School.

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Two Essays in Financial Econometrics

by

Yang YU

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

2013

This dissertation research explores two interesting problems in financial econometrics. In part one, we considers the problem of pricing European options in the presence of proportional transaction costs when the underlying stock price follows a jump-diffusion process. Based on utility maximization approach, the option pricing and hedging can be reformulated as a singular stochastic control problem. And furthermore, the value functions of the problem are the solutions of a free boundary problem, in particular, a partial integro-differential equation, under different boundary conditions. And we develop a coupled backward induction algorithm which is based on the connection between the free boundary problem and optimal stopping problem. And numerical examples are also provided. In part two, we focus on the dynamics of default risk with stochastic covariates in the presence of structural breaks. We consider a Cox type intensity model which is a classic model in survival analysis to deal with the counting process. Since it is widely used to to analyze the dynamics of default of firms to the effect of possible stochastic covariate processes. We assume there are multiple unknown structural breaks in the regression coefficients and we develop an estimation procedure for the regression coefficients and structural break points, which combines recent developments in estimating equations for counting process and inference on multiple structural breaks.

This thesis is dedicated to my wonderful mother LU, Yumei, who passed away in October 2010, with all loving memories. Without her unconditional and persistent love, I would have never become the individuals that I am today.

Mama:

No words can explain how much I miss you. Every day has been a struggle. It's been three years and it's still impossible for me to grasp that you are really gone. And days that if I allowed myself, I could cry and cry and cry. The sweetest moment for me is meeting you in my dream. You will be forever loved and missed. I had promised to make you proud and I hope that I have fulfilled that promise. I wish that you could still be here today to share with me the success of my graduation with a Doctor of Philosophy degree.

Table of Contents

List of Figures

List of Tables

Acknowledgements

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

I would like to express my deepest gratitude to my advisor, Professor Haipeng Xing, for exciting topics suggested, his excellent guidance, continuous support, caring and patience, providing me with an excellent atmosphere for doing research.

I would also like to thank Professor Wei Zhu, Professor Svetlozar Rachev and Professor Yixin Fang. It is my great honor to have them on my dissertation committee.

I would like to thank my wife Wei Zheng and my family. Without their encouragement, I would never have been able to finish my Ph.D. studies and without their love, all my achievement would mean nothing for me.

I also would like to thank all my groupmates and good friends, especially Ning Sun, Yifan Mo, Ying Cai and Lin Chen, for all their kindness, friendship, support and help.

Part I

European Option Pricing under Jump Diffusion with Proportional Transaction Costs

Chapter 1

Introduction

1.1 Stochastic Process

In probability theory, a *stochastic process* is a random process evolving with time. More precisely, a stochastic process is a collection of random variables X_t indexed by time. This is the probabilistic counterpart to a deterministic process (or deterministic system). The process could be in *discrete time* or *continuous time*. The random variables X_t will take values in a set which is called *state space*. The *state space* for a stochastic process can be *discrete*, i.e., a finite or countably infinite set, and be *continuous*, e.g., the real numbers $\mathbb{R}$ or d-dimensional space $\mathbb{R}^d$.

Definition 1.1.1. *Given a probability space $(\Omega, \mathcal{F}, P)$ and a measurable space (S, Σ), an S-valued stochastic process is a collection of S-valued random variables on Ω, indexed by a totally ordered set T ("time"). That is, a stochastic process X is a collection: $\{X_t : t \in T\}$ where each X_t is an S-valued random variable on Ω. The space S is then called the state space of the process.*

1.1.1 Brownian Motion

In finance, stochastic processes are used to model asset price fluctuations. Among them, the brightest star and the most fundamental process is Brownian Motion, which is a random process W_t with independent, stationary increments that follow a Gaussian distribution. Brownian Motion is the most widely and intensively studied stochastic process in finance and has been tied together with financial modelling from the very beginning of the latter, when Bachelier (1900) proposed to model the price S_t of an asset at the Paris Bourse as:

$$S_t = S_0 + \sigma W_t$$

The multiplicative version of Bachelier's model led to the commonly used Black-Scholes-Merton Model (see Merton 1973 and Black and Scholes 1973), where the log-price $\ln S_t$ follows a Brownian motion:

$$S_t = S_0 exp[\mu t + \sigma W_t]$$

or, in local form:

$$\frac{\mathrm{d}S_t}{S_t} = \sigma \mathrm{d}W_t + (\mu + \frac{\sigma^2}{2})\mathrm{d}t$$

The stock price process S_t is called *geometric Brownian motion.* Figure 1.1 shows two paths: the logarithm of the stock price for 3M Company between January 2005 and January 2010 (blue) and a sample path of Brownian motion, with the same average return and volatility as the stock over the five-year period considered (red). It may be difficult to tell which is which: the evolution of the stock does look like a Brownian motion process and that's why quantitative finance use Brownian motion to model price movements. However, for some trained eye, the difference between the two path is significant.

A very important property of Brownian motion is the *continuity* of its sample paths: a typical path $t \mapsto B_t$ is a continuous function of time. This property make us tell the two curves seen on Figure 1.1: unlike the Brownian motion, the 3M stock price underwent several abrupt upward and downward *jumps* during the period which appear as discontinuities in the price trajectory. Actually, even more striking than the comparison of price trajectories

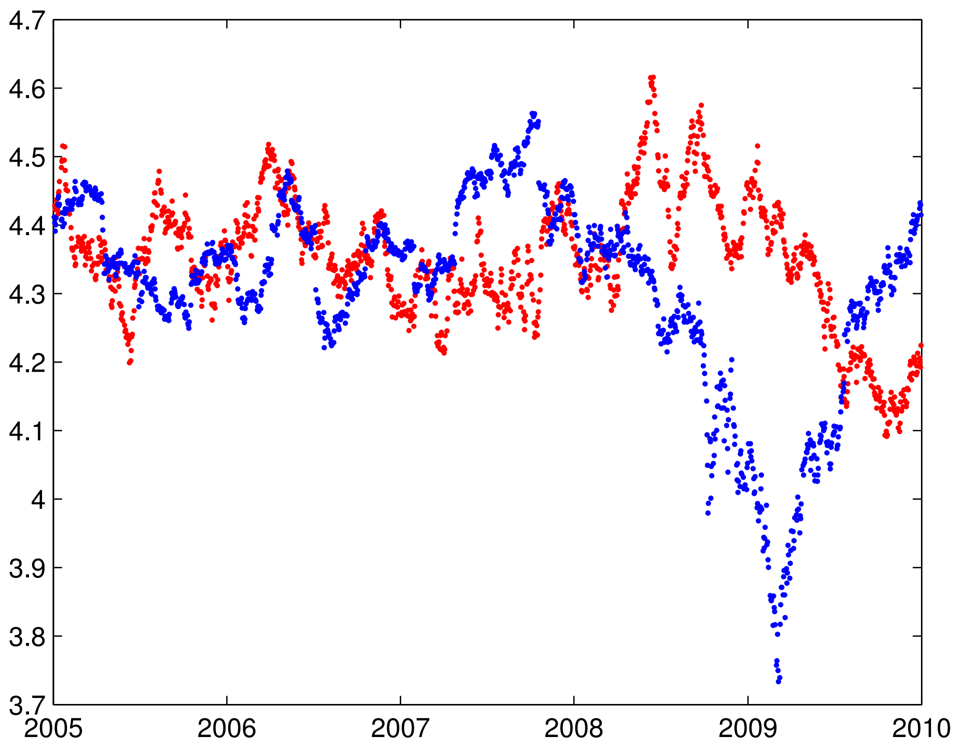

Figure 1.1: Log price for 3M from 2005 to 2010, compared with a sample path of Brownian motion with same annualized return (-0.0322) and volatility (25.69%). Can you tell which is which? (The red line plots the log price from simulation while the blue line plots the real log price of 3M.)

to those of Brownian path is the comparison of *returns*, i.e., increments of the log-price or called stock's log-return. Figure 1.2 compares the fifteen-year stock log-return on Microsoft to a Brownian motion with the same average return and volatility. While both the returns have the same variance, the returns from Brownian motion always have roughly the same amplitude whereas the real Microsoft Stock returns are widely dispersed in their amplitude and manifest frequent large peaks corresponding to “jumps” in the price. And this kind of high variability is a constantly observed feature of financial asset returns. In statistics, this results feature is called *heavy tails*: the tail of the distribution decays slowly at infinity and very large moves have a significant probability of occurring. However, Brownian motion or Gaussian model does not have this property. For example, for a normal distribution random variable, the probability of occurrence of a value six times the standard deviation is less than 10^{-8}. In a Gaussian model, a daily return of such magnitude occurs less than once in a million years! The Brownian motion model underestimates the risk in the market in a polite understatement. Even though we can generate diffusion process with arbitrary heavy tails by using some nonlinear diffusion process such as *local volatility* function proposed by Dupire (1994) and Derman and Kani (1994) and *stochastic volatility* models presented by Hull and White (1987) and Heston (1993) as follow:

$$\frac{\mathrm{d}S_t}{S_t} = \sigma(t, S_t)\mathrm{d}W_t + \mu\mathrm{d}t$$

we should know that they are not Gaussian processes. Furthermore, these heavy tails are produced at the price of obtaining highly varying (nonstationary) diffusion coefficients in local volatility function or unrealistically high values of “volatility of volatility” in diffusion-based stochastic volatility models. And even we can use the nonlinear diffusion process to generate heavy tails in returns, they *cannot* generate abrupt, large and discontinuous moves in prices. In a quest to incorporate the stylized properties of asset prices, we consider the *jump process* models.

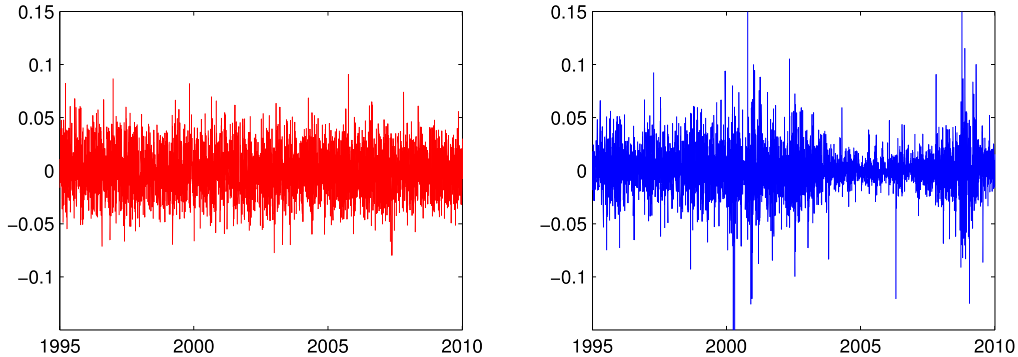

Figure 1.2: Log-return for stock of Microsoft from 1995 to 2010 (right), compared with a sample path of Brownian motion with same annualized return and volatility (left).

1.1.2 Jump Process

From the seminal paper Merton (1976) to the present date, various aspects of jump-diffusion models have been studied in the academic finance community (see Cont and Tankov, 2004, for a list of almost 400 references on the subject). In the last decade, more and more research departments of major banks started to accept jump-diffusion as a valuable tool in their day-to-day modeling. From the Table 1.1, we can clearly see the comparison between the diffusion models and the jump process models, which is coming from more than three decades research on financial modelling and risk management. And the brief comparison showed us that, beside having various empirical, computational and statistical features that have motivated their use in the first place, discontinuous models can provide *qualitatively different* information about the key issues of hedging, replication and risk.

In jump-diffusion model, Brownian motion is used to modulate the diffusion part, and for jump part, we always use *Poisson process* to build. Some statement about *Poisson process* is introduced below, and the proof can be found in Cont and Tankov (2004, Chapter 2)

Poisson process Take a sequence $\{\tau_i\}, i \geq 1$ of independent exponential random vari-

Table 1.1: Modelling market moves: diffusion models vs. models with jumps. See Cont. and Tankov (2004, Chapter 1)

Empirical facts	**Diffusion models**	**Models with jumps**
Large, sudden movements in prices.	Difficult: need very large volatilities.	Generic property.
Heavy tails.	Possible by choosing nonlinear volatility structure.	Generic property.
Options are risky investment.	Options can be hedged in a risk-free manner.	Perfect hedges do not exist: options are risky investment.
Markets are incomplete; some risks cannot be hedged.	Markets are complete.	Markets are incomplete .
Concentration: losses are concentrated in a few large downward moves.	Continuity: price movements are conditionally Gaussian; large sudden moves do not occur.	Discontinuity: jump / discontinuities in prices can give rise to large losses.
Some hedging strategies are better than others.	All hedging strategies lead to the zero residual risk, regardless of the risk measure used.	Hedging strategy is obtained by solving portfolio optimization problems.
Exotic options are hedged using vanilla (call/put) options.	Options are redundant: any payoff can be replicated by dynamic hedging with the underlying.	Options are not redundant: using vanilla options can allow to reduce hedging error.

ables with parameter λ, that is, with cumulative distribution function $P[\tau_i \geq y] = e^{-\lambda y}$ and let $T_n = \sum_{i=1}^{n} \tau_i$. The process

$$N_t = \sum_{n \geq 1} 1_{\{t \geq T_n\}}$$

is called the *Poisson process* with parameter λ. For example, if the waiting times between buses at a bus stop are exponentially distributed, the total number of buses arrived up to time t is a Poisson process. The trajectories of a Poisson process are piecewise constant (right-continuous with left limits or RCLL) and, as the same as Brownian motion, Poisson process has stationary and independent increment. That is, for every $t > s$ the increment $N_t - N_s$ is independent from the history of the process up to time s and has the same law as N_{t-s}.

Compound Poisson process In financial applications, it is of little interest to have a process with a single possible jump size. The *compound Poisson process* is a generalization where the waiting times between jumps are exponential but the jump size can have an arbitrary distribution: let N be a Poisson process with parameter λ and $\{Y_i\}, i \geq 1$ be a sequence of independent random variable with law f. The process

$$X_t = \sum_{i=1}^{N_t} Y_i$$

is called *compound Poisson process*. Its trajectories are RCLL and piecewise constant but the jump size are now random with law f. And it also has independent and stationary increment.

Jump-diffusions and Lévy processes Combining a Brownian motion with drift and a compound Poisson Poisson process, we can build up the simplest case of a jump-diffusion: a process which sometimes jumos and has a continuous but random evolution between the jump times:

$$X_t = \mu t + \sigma B_t + \sum_{i=1}^{N_t} Y_i \tag{1.1}$$

The well known model of this type in finance is the Merton model(see Merton, 1976). The process 1.1 is a Lévy process which is after the French mathematician Paul Lévy and the process has independent and stationary increments. In general, it consist three components: a deterministic and linear drift, a multiple of a Brownian motion and a pure jump process (see Rachev et al., 2005, Chapter 10). However, the class of Lévy process is not limited to jump-diffusion of the form 1.1. Lévy process could have infinitely many jumps in finite intervals which is called "infinite activity" (See the example provided by Madan and Seneta, 1990).

1.2 Stochastic Control

Stochastic control or optimal stochastic control is a subfield of control theory that deals with the existence of uncertainty either in observations of the data or in the things that drive the evolution of the data. It aims at predicting and minimizing the magnitudes and limits of the random deviations of a control system through optimizing the design of the controller. Such deviations occur when random noise and disturbance processes are present in a control system, so that the system does not follow its prescribed course but deviates from the latter by a randomly varying amount. The context may be either discrete time or continuous time. In contrast to deterministic signals, random signals cannot be described as given functions of time such as a step, a ramp, or a sine wave. The exact function is unknown to the system designer; only some of its average properties are known. A random signal may be generated by one of nature's processes, for instance, radar noise and wave-induced forces and moments on a radar antenna or a ship. Alternatively, it may be generated by human intelligence, for instance, the bearing of a zigzagging aircraft. Related research starts from 1950's and early 1960's. Some pioneers' works include Chernoff (1961), Bather and Chernoff (1966), Shiryaev (1961), Karatzas (1983) and Karatzas and Shreve (1985).

In very general terms, an optimal control problem consists of the following elements:

- *State process* $Z(\cdot)$. This process must capture of the minimal necessary information needed to describe the problem. For example, $Z(t) \in \mathcal{R}^d$ is influenced by the control and given the control process it has a Markovian structure. Usually its time dynamics is prescribed through an equation.

- *Control process* $\nu(\cdot)$ We need to describe the control set, U, in which $\nu(t)$ takes values in for every t. Applications dictate the choice of U. In addition to this simple restriction $\nu(t) \in U$, there could be additional constraints placed on control process. For instance, in the stochastic setting, we will require ν to be adapted to a certain filtration, to model the flow of information. Also we may require the state process to take values in a certain region (i.e., state constraint). This also places restrictions on the process $\nu(\cdot)$.

- *Admissible controls* $\mathcal{A}$ A control process satisfying the constraints is called an admissible control. The set of all admissible controls will be denoted by $\mathcal{A}$ and it may depend on the initial value of the state process.

- *Objective functional* $J(Z(\cdot), \nu(\cdot))$ This is the functional to be maximized (or minimized). Usually, J is given as an integral over time.

Then, the goal is to minimize (or maximize) the objective functional J over all admissible controls. The minimum value plays an important role in the analysis:

$$Value function := V = \inf_{\nu \in \mathcal{A}} J$$

The main problem in optimal control is to find the minimizing control process.

In finance context, the state variable in the stochastic differential equation is usually wealth or net worth, and the controls are the shares placed at each time in the various assets. Given the asset allocation chosen at any time, the determinants of the change in wealth are usually the stochastic returns to assets and the interest rate on the risk-free asset. The field of stochastic control has developed greatly since the 1970's, particularly in its applications to finance. Merton used stochastic control to study optimal portfolios of safe and risky assets. His work and that of Black-Scholes changed the nature of the finance literature.

1.3 Option Pricing Problem without Transaction Costs

Option pricing problem has been a popular and intensively discussed topic in finance since the publication of the Black and Scholes (1973) and Merton (1973). They gave a valuation for a *European call option*, a contract that confers on the buyer the right to buy at the *exercise time* T one share of a specified stock at an agreed price K (known as the strike price). Let S_t denote the stock price at time t. Apparently, the option is worthless if $S_T \leq K$, but it has positive value to the buyer and will be exercised if $S_T > K$. The writer ("seller") of the option has the obligation to deliver one share at time T for a cash payment of K if $S_T > K$. The pricing problem is to determine how much the writer should charge for issuing it at time t. It seems at first that the answer to this question must depend on the writer's attitude to risk and therefore that there can not be a "universal" pricing formula. However, Black and Scholes (1973) and Merton (1973) showed that, in certain circumstances, such a universal formula is indeed possible. Specifically, they assumed that the the price of risky asset ("stock"), S_t, is driven by a geometric Brownian motion, that a risk-free asset ("bond") with constant rate of return r, is available, and that funds may be transferred from bank to stock and vice versa without restrictions and transaction costs. There are no arbitrage opportunities in markets and security trading is continuous. Then it turns out that perfect hedging is possible and the value of an option equals the amount fo initial capital required for setting up the replicating portfolio.

1.3.1 Black-Scholes-Merton Model

In 1973, the Black-Scholes or Black-Scholes-Merton model was first published by Fischer Black and Myron Scholes in their paper, "The Pricing of Options and Corporate Liabilities", published in the Journal of Political Economy. They derived a stochastic partial differential equation, now called the Black-Scholes equation, which estimates the price of the option over time. The key idea behind the model is to hedge the option by buying and selling the underlying asset in just the right way, and consequently "eliminate risk". This hedge is called delta hedging and is the basis of more complicated hedging strategies such as those engaged in by investment banks and hedge funds. The hedge implies that there is a unique price for the option which does not depend on investor's preference and this is given by the Black-Scholes formula. Robert C. Merton was the first to publish a paper expanding the mathematical understanding of the options pricing model, and coined the term Black-Scholes options pricing model. Merton and Scholes received the 1997 Nobel Prize in Economics for their work.

The Black-Scholes-Merton model assumes the market consists of one risky asset, usually called the "stock", and one riskless asset, usually called the money market, cash, or bond.

Assumptions on the assets:

- The rate of return on the riskless asset is constant and thus called the risk-free interest rate.

- The instantaneous log returns of the stock price is an infinitesimal random walk with drift; more precisely, it is a geometric Brownian motion, and we will assume its drift and volatility is constant:

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W_t \tag{1.2}$$

- The stock does not pay a dividend.

Assumptions on the market:

- There is no arbitrage opportunity (i.e., there is no way to make a riskless profit).

- It is possible to borrow and lend any amount, even fractional, of cash at the riskless rate.

- It is possible to buy and sell any amount, even fractional, of the stock.

- The above transactions do not incur any fees or costs.

There are many different ways to deriving the Black-Scholes formula that follow use different types of mathematics, with different amounts of complexity and mathematical baggage. And even in the quantitative finance interview, the derivation of Black-Scholes formula has become an appetizer. Wilmott (2009) provided 12 different methods to derive the formula in his popular book. However, the original derivation of the Black-Scholes partial differential equation was via stochastic calculus, Ito's lemma and a simple hedging argument. (See Black and Scholes, 1973):

Assume that the underlying follows a geometry Brownian motion as 1.2. Use Π to denote the value of a portfolio of one long option position and a short position in some quantity Δ of the underlying:

$$\Pi = V(S, t) - \Delta S$$

The first term on the right is the option and the second term is the short asset position. The change in the portfolio value is due partly to the change in the option value and partly to the change in the underlying:

$$\mathrm{d}\Pi = \mathrm{d}V - \Delta \mathrm{d}S$$

From Ito's lemma, we can get:

$$\mathrm{d}\Pi = \frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial V}{\partial s}\mathrm{d}S + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\mathrm{d}t - \Delta \mathrm{d}S$$

The right-hand side of this contains two types of terms, the deterministic and the random. The deterministic terms are those with the dt, and the random terms are those with the dS. Pretending for the moment that we know V and its derivatives then we know everything about the right-hand side except for the value of dS, because this is random. But, these random terms can be eliminated by choosing

$$\Delta = \frac{\partial V}{\partial S}$$

After choosing the quantity Δ, we hold a portfolio whose value changes by the amount

$$\mathrm{d}\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\mathrm{d}t$$

This change is completely *riskless*. If we have a completely risk-free change dΠ in the portfolio value Π then it must be the same as the growth we would get if we put the equivalent amount of cash in a risk-free interest cash account:

$$\mathrm{d}\Pi = r\Pi \mathrm{d}t$$

Putting all of the above together to eliminate Π and Δ in favor of partial derivatives of V gives the Black-Scholes equation:

$$\mathrm{d}\Pi = \frac{\partial V}{\partial t} + rS\frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0$$

Solve this quite simple linear diffusion equation with the final condition

$$V(S,T) = max(S-K,0)$$

we will get the Black-Scholes call option formula:

$$C(S,t) = S_t N(d_1) - Ke^{-r(T-t)}N(d_2)$$

where

$$d_1 = \frac{\ln(S_t/K) + (r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \quad , \qquad d_2 = d_1 - \sigma\sqrt{T-t}$$

This derivation of the Black-Scholes equation is perhaps the most useful since it is readily generalizable (if not necessarily still analytically tractable) to different underlyings, more complicated models, and exotic contracts.

1.3.2 Kou's Double Exponential Jump-Diffusion Model

As we discussed in Section 1.1.2, more and more attention was paid to jump-diffusion model from 1970's, starting from seminal paper Merton (1976). In that work, Merton argued that the asset price fluctuations can be decomposed as the sum of "normal" vibrations, caused by temporary imbalance of supply and demand and other new information that causes marginal changes in the stock value, and “abnormal” vibrations, caused by the arrival of important new information which generates occasional and large impact on price. Furthermore, he modulated the normal and abnormal variations(jump) by a standard geometric Brownian motion and a Poisson process, respectively, and derived an option pricing formula. To capture asymmetric leptokurtic features in the underlying asset return distributions and volatility smiles in the option markets, Ait-Sahalia (2002), Carr and Wu (2003), Eraker et al (2003), Broadie et al (2007) developed models and methods that incorporate the occurrence of rare jumps in a price process that would otherwise follow a diffusion. Subsequent work has been moving towards considering other or more general Lévy process. For instance, Chan (1999) considered the problem of pricing contingent claims on a stock whose price process follows a geometric Lévy process, and Kou (2002) proposed a double exponential jump-diffusion model which can not only explain the two empirical phenomena but also provide rational expectations equilibrium framework and a psychological interpretation. An analytic solution was also obtained by the model. This double exponential jump-diffusion model successfully capture the leptokurtic feature of the return distribution and the “volatility smile” observed in option prices (see section 3 and section 5.3 in Kou, 2002). Andersen et al. (2002) demonstrate empirically that, for the S&P 500 data from 1980-1996, the normal jump-diffusion model has a much higher p-value (0.0152) than those of the stochastic volatility model (0.0008) and the Black-Scholes model ($< 10^5$). And in addition, the empirical tests performed in Ramezani and Zeng (1999) suggest that the double exponential jump-diffusion model fits stock data better than the normal jump-diffusion model. Therefore, the combination of the results in the two papers above gives us some empirical support of the double exponential jump-diffusion model. Therefore, we will use this model in our numerical studies(see Section 4.1), a short introduction be showed in this subsection, see Kou (2002) for more details.

The price of the underlying S_t is assumed to follow a double exponential jump-diffusion process:

$$\mathrm{d}S_t = \alpha S_{t-}\mathrm{d}t + \sigma S_{t-}\mathrm{d}W_t + S_{t-}\mathrm{d}\left(\sum_{i=1}^{N_t}(Q_i - 1)\right), \tag{1.3}$$

under the physical probability measure $\mathcal{P}$, where $\alpha > 0$ and $\sigma > 0$ are the expected return and diffusion volatility of the underlying asset, $\{W_t; t \geqslant 0\}$ is a standard Brownian motion with $W_0 = 0$, $\{N_t; t \geq 0\}$ is a Poisson process with rate λ, Q_i is a sequence of independent and identically distributed positive random variables such that the jump $Y = \log Q$ has asymmetric double exponential distribution with density:

$$f_Y(y) = \begin{cases} p\eta_1\, e^{-\eta_1 y} & \text{if } y \geq 0, \\ (1-p)\eta_2\, e^{\eta_2 y} & \text{if } y < 0. \end{cases}$$

Here $0 \leq p \leq 1$, and p and $1-p$ represent the probability of positive and negative jumps, respectively. The parameters η_1 and η_2 are assumed to satisfy $\eta_1 > 1$ and $\eta_2 > 0$.
Alternatively, under the risk-neutral measure $\mathcal{P}^*$,

$$\frac{\mathrm{d}S_t}{S_{t-}} = (r - \lambda\psi)\mathrm{d}t + \sigma\mathrm{d}W_t + \mathrm{d}\left(\sum_{i=1}^{N_t}(Q_i - 1)\right), \tag{1.4}$$

where $\psi = \mathbb{E}(Q-1)$. We can get the unique solution of the equation above:

$$S_t = S_0 exp\left[\left(r - \frac{\sigma^2}{2} - \lambda\psi\right)t + \sigma W_t\right]\prod_{i=1}^{N_t} Q_i$$

To price a European option in the jump-diffusion model, it remains to compute the expectation, under the measure $\mathcal{P}^*$, of the discounted final payoff of the option. In particular, the price of a European call option at time t is given by

$$\begin{aligned} C_t &= \mathbb{E}_*[e^{-r(T-t)}(S_T - K)_+] \\ &= \mathbb{E}_*\left[e^{-r(T-t)}\left(S_t exp[(r - \frac{\sigma^2}{2} - \lambda\psi)(T-t) + \sigma\sqrt{T-t}\epsilon]\prod_{i=1}^{N_T} Q_i - K\right)_+\right] \end{aligned}$$

where T is the expiration time, $(T-t)$ is the time to expiration measure in years, K is the strike price, $(x)_+ = max(0, x)$ and ϵ is a standard normal random variable. Kou (2002) showed that c_t is analytically tractable as

$$\begin{aligned} C_t &= \sum_{n=1}^{\infty}\sum_{j=1}^{n} e^{-\lambda(T-t)}\frac{\lambda^n(T-t)^n}{n!}\frac{2^j}{2^{2n-1}}\binom{2n-j-1}{n-1}(A_{1,n,j} + A_{2,n,j} + A_{3,n,j}) \\ &\quad + e^{-\lambda(T-t)}[S_t e^{-\lambda\psi(T-t)}\Phi(h_+) - Ke^{-r(T-t)}\Phi(h_-)] \end{aligned} \quad (1.5)$$

where $\Phi(\cdot)$ is the CDF of the standard normal random variable,

$$
\begin{aligned}
A_{1,n,j} &= S_t e^{-\lambda\psi(T-t)+n\kappa}\frac{1}{2}\left(\frac{1}{(1-\eta)^j}+\frac{1}{(1+\eta)^j}\right)\Phi(b_+) - e^{-r(T-t)}K\Phi(b_-),\\
A_{2,n,j} &= \frac{1}{2}e^{-r(T-t)-\omega/\eta+\sigma^2(T-t)/(2\eta^2)}K\\
&\quad\times\sum_{i=0}^{j-1}\left(\frac{1}{(1-\eta)^{j-i}}-1\right)\left(\frac{\sigma\sqrt{T-t}}{\eta}\right)^i\frac{1}{\sqrt{2\pi}}Hh_i(c_-),\\
A_{3,n,j} &= \frac{1}{2}e^{-r(T-t)+\omega/\eta+\sigma^2(T-t)/(2\eta^2)}K\\
&\quad\times\sum_{i=0}^{j-1}\left(1-\frac{1}{(1+\eta)^{j-i}}\right)\left(\frac{\sigma\sqrt{T-t}}{\eta}\right)^i\frac{1}{\sqrt{2\pi}}Hh_i(c_+),\\
b_{\pm} &= \frac{\ln(S_t/K)+(r\pm\sigma^2/2-\lambda\psi)(T-t)+n\kappa}{\sigma\sqrt{T-t}},\\
h_{\pm} &= \frac{\ln(S_t/K)+(r\pm\sigma^2/2-\lambda\psi)(T-t)}{\sigma\sqrt{T-t}},\\
c_{\pm} &= \frac{\sigma\sqrt{T-t}}{\eta}\pm\frac{\omega}{\sigma\sqrt{T-t}}\\
\omega &= \ln(K/S_t)+\lambda\psi(T-t)-(r-\sigma^2/2)(T-t)-n\kappa,\\
\psi &= \frac{e^\kappa}{1-\eta^2}-1,\\
\kappa &= \mathbb{E}(\log(Q))
\end{aligned}
$$

and the $Hh(\cdot)$ function are defined as

$$
Hh_n(x) = \frac{1}{n!}\int_x^\infty (s-x)^n e^{-s^2/2}\mathrm{d}s, \qquad n = 0, 1, \ldots,
$$

and $Hh_{-1}(x) = exp(-x^2/2)$, which is $\sqrt{2\pi}f(x)$ with $f(x)$ being the probability density function of a standard normal random variable. The $Hh_n(x)$ function satisfy the recursion:

$$nHh_n(x) = Hh_{n-2}(x) - xHh_{n-1}(x), \qquad n \geq 1,$$

with initial values $Hh_{-1} = e^{-x^2/2}$ and $Hh_0(x) = \sqrt{2\pi}\Phi(-x)$.

This pricing formula involves an infinite series, but its numerical value can be approximated quickly and accurately through truncation (e.g. the first 10 terms). Also, if $\lambda = 0$ which means there is no jumps, then it is easily seen that C_t in 1.5 reduces to the Black-Scholes formula for a call option discussed in last subsection. And through *Put-Call Parity*, we can get the put option price as:

$$P_t = C_t + Ke^{-r(T-t)} - S_t$$

1.4 Option Pricing Problem with Transaction Costs for Diffusion Process

No transaction costs in the replicating is an important assumption in Black-Scholes-Merton model. However, We do not live in a Black-Scholes world. In real world, there are, of course, costs in trading. Especially in the presence of transaction costs proportional to the amount of trading, such a continuous trading strategy which is required by Black-Scholes "delta-hedging" portfolio becomes impractical and prohibitively expensive. Hence it is impossible to perfectly replicate the option in this setting when there are transaction costs.

To deal with the problem of option pricing and hedging with transaction costs, several approaches had been proposed for the case that the underlying stock price follows a geometric

Brownian motion. One approach is based on super-replication in a discrete-time setting and tries to find trading strategies which produce payoffs at expiration that are larger than or equal to the option payoff. In these works, Leland (1985) proposed an adjustment in volatility which is used in the Black-Scholes delta to get the option payoff at expiration inclusive of transaction costs. However, this modified strategy is not *self-financing*. After that, Bole and Vorst (1992) still worked in discrete-time setting (binomial tree)and construct a *self-financing* replicating strategy. Explicit portfolio weights at each node of the binomial tree can be computed by a backward induction procedure. However, the method need the user to specify a revision interval and how to do so optimally is not clear. Soner et al. (1995) proved that when the width of the revision interval approaches 0, the cost of the option approaches the price of a single share of underlying stock. So the trivial strategy of buying one share of underlying stock and holding to maturity is the least expensive way of super-replicating the option in a continuous-time model with transaction costs. Bensaid et al (1992), through the binomial tree model, derived bounds on the option value by minimizing the initial cost of the *self-financing* strategy used to yield a super-replicating portfolio of stock and bond at expiration. And they pointed out that, by rebalancing only in the earlier periods, it is possible to have a super-replicating portfolio that is less expensive than the corresponding replicating portfolio. In general, the optimal discrete-time super-replicating strategy is such that the investor with an option position does not transact at a trading date if the inherited amount of stock is in a certain range (which depends on the past history of the stock price); otherwise he adjusts his portfolio back to this range. However, the cost minimization problem associated with super-replication is path dependent and that the dynamic programming algorithm is computationally expensive if the number of periods is not sufficiently small. Edirisinghe et al. (1993) developed a linear programming algorithm and a two-stage dynamic programming method to approximate the optimal solution. More recently, Primbs (2009) provided an alternative formulation of super-replication in terms of the first two moments of the replication error.

Another approach examines the difference between the desired payoff at maturity and the realized cash flow from a hedging strategy, and tries to achieve the best possible trade off between the cost of payoff and the risk. A pioneer work along this line is Hodges and Neuberger (1989), who formulated the problem of option pricing and hedging by maximizing the investor's expected utility of terminal wealth. Using an indifference argument, the selling or buying price of an option is defined as the amount of money that would make an investor indifferent, in terms of expected utility, between trading in the market with and without a position in the option. They interpreted the "hedging" of the option as the difference in the two trading strategies, with and without the option. This approach involves the value functions of singular stochastic control problems. The nature of optimal hedging is that an investor with an option position rebalances his portfolio only when the number of the stock shares falls out the lines ("too much" line and "too few" line). Davis et al. (1993) modified certain settings of Hodges and Neuberger (1989) and developed rigorously the model of Hodges and Neuberger (1989) for a market with proportional transaction costs. In particular, Davis et al. (1993), Clewlow and Hodges (1997)and Zakamouline (2006) showed that the option pricing problem with transaction costs involves solving two singular stochastic control problems formulated by Davis and Norman (1990), and developed, for the negative exponential utility function, numerical algorithms to compute the optimal hedge and option price by making use of discrete-time dynamic programming for an approximating binomial tree for the stock price. Further contributions to the study of the utility based option pricing approach and numerical methods with proportional transaction costs include Whalley and Wilmott (1997) which developed asymptotic approximations for these hedging strategies and option prices as the transaction costs approach 0, Constantinides and Zariphopoulou (1999,2001) which provided option price bounds under general utility functions rather than the negative exponential utility function commonly adopted for numerical studies, and Andersen and Damgaard (1999) consider more than one risky security, and some others. In this subsection, we will briefly introduce model of Davis et al. (1993). Even though the computational method presented by Davis et al. (1993) can not efficiently and accurately solve the problem involving the jump-diffusion model, the utility maximization is still the key idea and basic frame in our whole work.

In the model of Davis et al. (1993), the price of European option is defined in terms of a utility maximization problem. In short, the option price, which is to the writer, is obtained by a comparison of the maximum utilities of trading with and without the obligation of fulfilling the option contract at expiry. The asset price S_t is assumed to follow the *geometry Brownian motion*:

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W$$

where μ and σ are constant and W si a Brownian motion.

When the utility function takes the special form $\mathcal{U}(x) = 1 - \exp(-\gamma x)$ in which γ is the risk aversion parameter, Davis et al. (1993) found that the option price $V(S,t)$ is given by

$$V(S,t) = \frac{\delta(T,t)}{\gamma} \log\left(\frac{Q_\omega(S,0,t)}{Q_1(S,0,t)}\right)$$

where T is te expiry date, $\delta(T-t) = e^{-r(T-t)}$, and $Q_1(S,x,t)$ and $Q_\omega(S,x,t)$ both satisfy the following equation

$$\min\left(\frac{\partial Q}{\partial x} + \frac{\gamma(1+\epsilon)SQ}{\delta}, -\frac{\partial Q}{x} - \frac{\gamma(1-\epsilon)SQ}{\delta}, \frac{\partial Q}{t} + \mu S \frac{\partial Q}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 Q}{\partial S^2}\right) = 0$$

Here ϵ is the transaction costs: A trade of N shares will result in a loss of ϵNS. This cost structure represents bid-off spread, or more generally commissions and costs that are proportional to the value of the assets traded. In Davis et al. (1993), they consider the slightly more general case in which there are different levels of cost for buying and selling. The independent variable x measures the number of shares held in the portfolio and maybe changed according to optimally hedging transaction. The two functions Q_1 and Q_ω must satisfy certain final conditions, analogous to the payoff profile of the option; for example, for a European call option

$$Q_1(S, x, T) = \exp(-\gamma c(S, x))$$

and

$$Q_\omega(S, x, T) = \begin{cases} \exp(-\gamma c(S, x)) & S \leq K \\ \exp(-\gamma(c(S, x) + K - S)) & S > K \end{cases}$$

where

$$c(S, x) = \begin{cases} (1+\epsilon)yS & y < 0 \\ (1-\epsilon)yS & y \geq 0 \end{cases}$$

So the final condition for the second problem (Q_ω) is equal to that of the first problem (Q_1) modified by the effects of the potential liability at expiry of the European call (after transaction costs). Note we are assuming here that the option is cash settled. For options with delivery of the asset on exercise the below remains the same and the final condition merely alter.

Then this stochastic control problem can be expressed as a free boundary problem. And they explained that the (S, x) space divided into three regions, shown schematically in Figure 1.3. The writer, or issuer, of the option must always maintain the portfolio in the region of the (S, x) space bounded by the two outer curves, while inside this region he does not transact. If the number of shares of his portfolio hits the top boundary, he must sell shares; if the number of shares of his portfolio hits the bottom boundary, he must buy shares. That means when the portfolio goes to the edge of this *no-transaction region*, he must trade so as to just stay inside. And the simple analytical expressions for all three of these curves is not

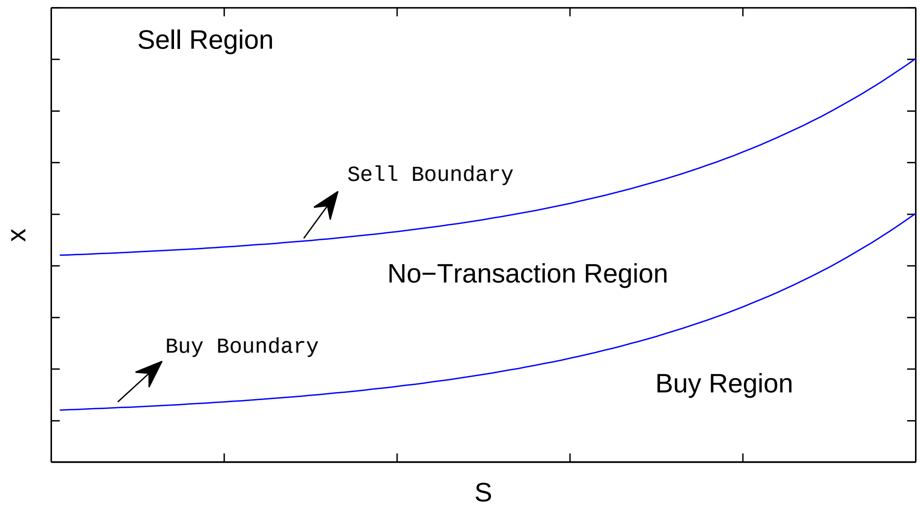

Figure 1.3: A schematic diagram of (S, x) space showing the buy, sell and no-transaction regions

hard to derived.

In the *buy region* we have

$$\frac{\partial Q}{\partial x} + \frac{\gamma(1+\epsilon)SQ}{\delta} = 0$$

In the *sell region* we have

$$\frac{\partial Q}{x} + \frac{\gamma(1-\epsilon)SQ}{\delta} = 0$$

In the *no-transaction region* we have

$$\frac{\partial Q}{t} + \mu S \frac{\partial Q}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 Q}{\partial S^2} = 0$$

To solve the option pricing problem under diffusions, Kushner and Dupuis (1992), Davis et al. (1993) and Zakamouline (2006) have developed a numerical scheme that involves a consistent Markov chain approximation of continuous-time price processes and then solved an appropriate optimization problem by discrete time dynamic programming. In particular, their scheme is based on weak convergence of probability measures and uses Markov chain approximation and discrete time dynamic programming algorithm. In their scheme, the discrete state is $\mathbb{X}(t, S, x, y)$, whose elements denote time, stock price, number of shares and amount in the bank in a discrete space. The value functions Q_1 and Q_w are given a value at the final time by using the boundary conditions for the continuous value functions over the discrete subspace (S, x, y), and then they are estimated by proceeding backward in time by using the discrete time algorithm. As in the continuous time case, this algorithm is the same for both value functions and is derived below for a value function denoted by $Q^\rho(t, S, x, y)$, where ρ is a discretization parameter, which depends on the discrete time interval δt. If δt and the resolution of the η axis $\delta\eta$ are sent to zero, then the above discrete value function converges to a viscosity subsolution and a viscosity supersolution of the PDE. Therefore, all the discrete value functions converge to their continuous counterparts; this is due to the uniqueness of the viscosity solution. Though the extension of the scheme from diffusion-only to jump diffusion processes is possible, it is computationally expensive as the solutions of the singular stochastic control problem (2.22) require the determination of both when to apply control and how much control to apply. So we design new method called *coupled backward induction algorithm* that can substantially reduce computational complexities. The key idea of the proposed algorithm is recently introduced by Lai and Lim (2009), and makes use of the connection of (2.22) to an optimal stopping problem. Some details will be showed in Section 3.3

1.5 Outline

This dissertation research is motivated by the concerns mentioned above. In real financial market, there is, of course, cost in each buying or selling transaction. The assumption of no transaction cost in Black-Scholes-Merton Model is impractical and when we want to consider the option pricing problem in real world, the first assumption we want to break is on the transaction cost. In another way, many research work have show the disadvantage of diffusion process setting for the underlying asset price process. Based on empirical study in the last decade, we think the jump-diffusion model is a much better framework setting for the price process compared to the diffusion process. So, to sum it up, in this work, we relax both of the two impractical assumptions to studies the problem of option pricing under general jump-diffusion processes with proportional transaction costs by generalizing the utility-based approach developed by Davis et al. (1993) and proposing a coupled backward induction algorithm to compute the solutions. Specially, under the assumption that the underlying stock price follows a geometric Lévy process, we formulate the problem of option pricing with proportional transaction costs as the maximization of the investor's expected utility of terminal wealth, and demonstrate that the implied singular stochastic control problem can be reduced to a free boundary problem for a partial integro-differential equation. We then develop a numerical algorithms, which is called coupled backward induction algorithm, to solve the equations for the negative exponential utility function and compute the buy-sell boundaries and value functions of the maximization problem simultaneously, which is different from Davis et al. (1993), Clewlow and Hodges (1997) and Zakamouline (2006) who used discrete-time dynamic programming for an approximating binomial tree for the stock price.

The singular stochastic control problem is introduced in Chapter 2: Section 2.1 introduces option pricing problem under utility maximization; Section 2.2 PIDE by HJB Equation. Section 2.3 Solutions for the negative exponential utility function. In Chapter 3, we provide a numerical method which is called "coupled-backward induction algorithm" to solve the PIDE. Chapter 4 provides intensive simulation studies that investigate the impact of jump component in the stock price process and transaction costs on option price and the implied hedging costs. Some concluding remarks are given in Chapter 5.

Chapter 2

A Singular Stochastic Control Problem

Options contracts have been known for many centuries and the trading activity increased since 1973 when options were issued with standardized terms and traded through a guaranteed clearing house at the Chicago Board Options Exchange. Today many options are created in a standardized form and traded through clearing houses on regulated options exchanges, while other over-the-counter options are written as bilateral, customized contracts between a single buyer and seller, one or both of which may be a dealer or market-maker. Options are part of a larger class of financial instruments known as derivative products, or simply, derivatives. In definition, an European option is a contract which gives the buyer (the owner) the right, but not the obligation, to buy or sell an underlying asset or instrument at a specified price(called "strike price K") on a specified date(called "expiration date T"). The option seller or writer incurs a corresponding obligation to fulfill the transaction which is to sell or to buy, if the owner elects to "exercise" the option prior to expiration. The buyer pays a premium to the seller for this right. An option which conveys to the owner the right to buy something at a specific price is referred to as a call; an option which conveys the right of the owner to sell something at a specific price is referred to as a put. Both are commonly traded, but for clarity, we only discussed the call option.

For a European call option, a contract that confers on the buyer the right to buy at the *expiration date* T one share of a specified stock at strike price K, let S_t denote the stock price at time t. Apparently, the option is worthless if $S_T \leq K$, but it has positive value to the buyer and will be exercised if $S_T > K$. The writer ("seller") of the option has the obligation to deliver one share at time T for a cash payment of K if $S_T > K$. The pricing problem is to determine how much the writer should charge for issuing it at time t.

2.1 Option Pricing Problem based on Utility Maximization

Suppose that an investor is provided with an opportunity to enter into a position in a European call option written on a stock with expiration date T and strike price K. The price of the stock is assumed to follow a geometric Lévy process

$$\mathrm{d}S_t = \alpha S_{t-}\mathrm{d}t + \sigma S_{t-}\mathrm{d}W_t + S_{t-}\int_{-1}^{\infty} \eta \widetilde{N}(\mathrm{d}t, \mathrm{d}\eta). \tag{2.1}$$

Here the mean rate of the stock return $\alpha > 0$ and the volatility $\sigma > 0$ are constants, and $W_t, t \geq 0$ is a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ with $W_0 = 0$. The Poisson random measure $\widetilde{N}$ is $\mathcal{F}_t$-centered, that is,

$$\widetilde{N}(t, A) = N(t, A) - E[N(t, A)] = N(t, A) - t\nu(A),$$

in which Poisson random measure $N(t, A)$ measures the number of jumps with amplitude in $A \subset (-1, \infty)$ up to and including time t, N has a time-homogeneous intensity, $E[N(t, A)] = t\nu(A)$, and ν is the Lévy measure associated to N. Note that to remain the price process $S_t > 0$ for all $t \geq 0$, we only allow jump sizes $\eta > -1$. We also assume that, for technical convenience,

$$\int_{-1}^{\infty} 1 \vee \eta^2 d\nu(\eta) < \infty.$$

In the presence of proportional transaction costs, the investor pays $0 < \zeta < 1$ and $0 < \mu < 1$ of the dollar value transacted on purchase and sale of the underlying stock, respectively. Let x_t denote the number of shares held in stock and y_t the dollar value of the investment in bond which pays a fixed risk-free rate r. The investor's position (x_t, y_t) in stock and bond is driven by:

$$\mathrm{d}x_t = \mathrm{d}L_t - \mathrm{d}M_t, \tag{2.2}$$

$$\mathrm{d}y_t = ry_t\mathrm{d}t - aS_t\mathrm{d}L_t + bS_t\mathrm{d}M_t, \tag{2.3}$$

with $a = 1 + \zeta$ and $b = 1 - \mu$, where L_t and M_t are nondecreasing and non-anticipating processes and represent the cumulative number of shares of stock bought or sold, respectively, within the time interval $[0, t]$, $0 \le t \le T$. Equations (2.1), (2.2) and (2.3) composes the market model in the time interval $[0, T]$, which describes a degenerate jump diffusion process in $\mathbb{R}^3$.

Consider a class of trading strategies such that L_t and M_t are absolutely continuous processes, given by

$$L_t = \int_0^t l_u \mathrm{d}u \qquad \text{and} \qquad M_t = \int_0^t m_u \mathrm{d}u, \tag{2.4}$$

where l_u and m_u are positive and uniformly bounded by $\xi < \infty$. Then equations (2.2) and (2.3) can be rewritten as:

$$\mathrm{d}x_t = l_t\mathrm{d}t - m_t\mathrm{d}t \tag{2.5}$$

$$\mathrm{d}y_t = ry_t\mathrm{d}t - aS_tl_t\mathrm{d}t + bS_tm_t\mathrm{d}t \tag{2.6}$$

As a special case, consider trading strategies which make an instantaneous change in the amount of shares of stock whenever the stock position at time t moves outside some allowable range $[X_b(t,S_t), X_s(t,S_t)]$:

$$\mathrm{d}L_t = (X_b(t,S_t) - x_{t-})^+ \qquad \text{and} \qquad \mathrm{d}M_t = (X_s(t,S_t) - x_{t-})^+$$

The corresponding stock position is given by

$$x_t = \begin{cases} X_b(t,S_t) & \text{if} \quad x_{t-} < X_b(t,S_t), \\ X_s(t,S_t) & \text{if} \quad x_{t-} > X_s(t,S_t), \\ x_{t-} & \text{if} \quad X_b(t,S_t) \leqslant x_t \leqslant X_s(t,S_t). \end{cases}$$

Denote the terminal settlement value of the stock and option by $Z^i(S_T, x_T)$, where $i = 0$, s and b indicates the investor's position in the option: no call option, short call, and long call, respectively. When there is no option position, or $i = 0$, the liquidated value of stock is given by

$$Z^0(S,x) = xS(a\mathbb{I}_{\{x<0\}} + b\mathbb{I}_{\{x\geqslant 0\}}). \tag{2.7}$$

Since when you have short position in stock($x < 0$), you need to buy the shares back from market and pay ζ of dollar value transacted on purchase. In contrast, when you have long position in stock($x > 0$), you need to liquidated all shares and pay μ of dollar value transacted on sale.

If the option is *cash settled*, the option writer delivers $(S_T - K)^+$ in cash at T, so

$$Z^i(S,x) = Z^0(S,x) - (S-K)\Delta^i(S), \qquad i = s, b, \tag{2.8}$$

where $\Delta^s(S) = \mathbb{I}_{\{S>K\}}$ (short call) and $\Delta^b(S) = -\mathbb{I}_{\{S>K\}}$ (long call). If the option is *asset* *settled*, then the writer delivers one share of stock in return for a payment of K when the buyer exercises the option at maturity T, so

$$Z^i(S,x) = Z^0(S, x - \Delta^i(S)) + K\Delta^i(S), \qquad i = s, b. \tag{2.9}$$

Note that the cases of (2.8)-(2.9) implies the trade of 0 share of stock for $i = s$, and b with cash settlement, and the trade of $\Delta^i(S_T)$ share of stock for $i = s$ and b with asset settlement. Using the self-financing argument of Bensaid et al. (1992) and assuming that the investor can choose any time in $[t, T]$ to trade and the first trade after time t occurs at time τ, we then have $x_u = x_t$ for all $u \in [t, \tau)$, and the dollar value of the investment at time τ is given by

$$y_\tau = y_t\, e^{r(\tau - t)} - S_\tau(a\mathrm{d}L_\tau - b\mathrm{d}M_\tau),$$

in which $S_\tau(a\mathrm{d}L_\tau - b\mathrm{d}M_\tau)$ is the cost of the first trade. In general, suppose trades after time $\tau_0 := t$ occur at times $\tau_1 < \tau_2 < \cdots < \tau_n$, we have

$$y_T = y_t\, e^{r(T-t)} - \sum_{i=0}^{n} e^{r(T-\tau_i)} S_{\tau_i}(a\mathrm{d}L_{\tau_i} - b\mathrm{d}M_{\tau_i}).$$

Extending the argument to continuous time and continuous states yields

$$y_T = y_t\, e^{r(T-t)} - \Psi(L, M; t, T) \tag{2.10}$$

where

$$\Psi(L, M; t, T) = a \int_{[t,T)} e^{r(T-u)} S_u \mathrm{d}L_u - b \int_{[t,T)} e^{r(T-u)} S_u \mathrm{d}M_u \tag{2.11}$$

is the total trading cost incurred over $[t, T)$. In the mean while, the total hedging cost incurred in $[t, T]$ is expressed as

$$C^i(L, M; t, T) = \Psi(L, M; t, T) - Z^i(S_T, x_T), \qquad i = 0,\ s,\ b. \tag{2.12}$$

Therefore, combining (2.10)-(2.12) yields the investor's terminal wealth in terms of the total hedging cost

$$y_T + Z^i(S_T, x_T) = y_t\, e^{r(T-t)} - C^i(L, M; t, T), \qquad i = 0,\ s,\ b,$$

Suppose that the writer's utility $U : \mathbb{R} \to \mathbb{R}$ is a concave and increasing function with $U(0) = 0$. We assume that the investor's goal is to maximize the expected utility of terminal wealth under the market model (2.1)-(2.3)

$$V^i(t, S, x, y) = \sup_{L,M} \mathbb{E}[U(y_T + Z^i(S_T, x_T))|S_t = S, x_t = x, y_t = y], \tag{2.13}$$

which corresponds to no call option ($i = 0$), short call ($i = s$) and long call ($i = b$). With the given utility function (2.13), the option price can be derived from the indifference argument which is similar to utility equivalence pricing principle in economics. In particular, the writing price of an option is defined as the amount of money that makes the investor indifferent, in terms of expected utility, between entering into the market with and without writing the option. At this *reservation* price, the investor is indifferent between selling (or buying) an option and doing nothing. Denote the reservation price of selling (or buying) an option as the amount of cash P^s (or P^b) required initially to provided the same expected utility as not enter into this position to the investor, P^s and P^b satisfy the following equations:

$$V^s(0, S, x, y + P^s(S, x)) = V^0(0, S, x, y) = V^b(0, S, x, y - P^b(S, x)). \tag{2.14}$$

2.2 Partial Intergro Differential Equation through HJB Equation

Now, taking equations (2.1)-(2.3) as a vector stochastic differential equation with controlled drift, we consider a class of trading strategies as absolutely continuous process, as given by (2.4) and the stochastic control problems for the utility maximization problem (2.13). Now, we can derive the Hamilton-Jacobi-Bellman equations for the value function V^i as follow:

$$\begin{aligned} V(t,S_t,x_t,y_t) &= \sup_{(l_t,m_t)} \mathbb{E}\left(V(t+\mathrm{d}t, S_{t+\mathrm{d}t}, x_{t+\mathrm{d}t}, y_{t+\mathrm{d}t})|\mathcal{F}_t\right) \\ &= \sup_{(l_t,m_t)} \mathbb{E}\Bigg(V(t,S_t,x_t,y_t) + \frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial V}{\partial y}\mathrm{d}y + \frac{\partial V}{\partial x}\mathrm{d}x \\ &\quad + \frac{\partial V}{\partial S}\mathrm{d}S + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(\mathrm{d}S)^2|\mathcal{F}_t\Bigg) \end{aligned}$$

Distract $V(t,S_t,x_t,y_t)$ from both sides, we can get

$$\begin{aligned} 0 &= \sup_{(l_t,m_t)} \mathbb{E}\Bigg(\frac{\partial V}{\partial t}\mathrm{d}t + \frac{\partial V}{\partial y}(ry_t - aS_t l_t + bS_t m_t)\mathrm{d}t + \frac{\partial V}{\partial x}(l_t - m_t)\mathrm{d}t \\ &\quad + \frac{\partial V}{\partial S}S(r\mathrm{d}t + \sigma \mathrm{d}W_t + \int_{-1}^{\infty} \eta \widetilde{N}(\mathrm{d}t,\mathrm{d}\eta)) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\mathrm{d}t\Bigg) \end{aligned}$$

And then,

$$\begin{aligned} \Rightarrow 0 &= \max_{l_t,m_t}\left((\frac{\partial V}{\partial x} - aS_t\frac{\partial V}{\partial y})l_t - (\frac{\partial V}{\partial x} - bS_t\frac{\partial V}{\partial y})m_t\right) + \frac{\partial V}{\partial t} + ry\frac{\partial V}{\partial y} \\ &\quad + \alpha S\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \\ &\quad + \int_{-1}^{\infty}\Big[V(t,S(1+\eta),x,y) - V(t,S,x,y) - \eta S_t\frac{\partial V}{\partial S}\Big]d\nu(\eta). \end{aligned}$$

From the equation above, we can get:

	I	II	III	IV
$\frac{\partial V}{\partial x} - aS_t\frac{\partial V}{\partial y}$	$\geqslant 0$	< 0	$\leqslant 0$	$\geqslant 0$
$\frac{\partial V}{\partial x} - bS_t\frac{\partial V}{\partial y}$	> 0	$\leqslant 0$	$\geqslant 0$	$\leqslant 0$

For situation I: $m_t = 0$ and $l_t = c$, which means it is in *Buy region*

For situation II: $m_t = c$ and $l_t = 0$, which means it is in *Sell region*

For situation III: $m_t = 0$ and $l_t = 0$, which means it is in *No transaction region*

Situation IV is impossible because the value function are increasing functions of x and y.

(I) In the buy region, the value function remains constant along the path of the state, dictated by the optimal trading strategy, and therefore,

$$\begin{aligned} V(t,S,x,y) &= V(t,S,x+\mathrm{d}x, y-aS\mathrm{d}x) \\ &= V(t,S,x,y) + \frac{\partial V}{\partial x}\mathrm{d}x - \frac{\partial V}{\partial y}aS\mathrm{d}x \end{aligned}$$

And then,

$$\Rightarrow \frac{\partial V}{\partial x} - aS\frac{\partial V}{\partial y} = 0 \qquad \text{Buy region boundary}$$

(II) In the sell region, similarly, the value function obeys the following equation:

$$\begin{aligned} V(t,S,x,y) &= V(t,S,x-\mathrm{d}x, y+bS\mathrm{d}x) \\ &= V(t,S,x,y) - \frac{\partial V}{\partial x}\mathrm{d}x + \frac{\partial V}{\partial y}bS\mathrm{d}x \end{aligned}$$

becomes,

$$\Rightarrow \frac{\partial V}{\partial x} - bS\frac{\partial V}{\partial y} = 0 \qquad\qquad \text{Sell region boundary}$$

(III) In the no transaction region, the value function obeys the same set of equations obtained for the class of absolutely continuous trading strategies, and therefore the Hamilton-Jacobi-Bellman is given by:

$$\max_{l_t, m_t}\Big\{\left(\frac{\partial V^i}{\partial x} - aS_t\frac{\partial V^i}{\partial y}\right)l_t - \left(\frac{\partial V^i}{\partial x} - bS_t\frac{\partial V^i}{\partial y}\right)m_t\Big\} + \mathcal{L}_1 V^i = 0, \tag{2.15}$$

for $(t, S_t, x_t, y_t) \in [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$, in which the operator $\mathcal{L}_1$ is defined as

$$\begin{aligned}\mathcal{L}_1\phi := \frac{\partial \phi}{\partial t} + ry\frac{\partial \phi}{\partial y} + \alpha S\frac{\partial \phi}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 \phi}{\partial S^2} \\ + \int_{-1}^{\infty}\Big[\phi(t, S(1+\eta), x, y) - \phi(t, S, x, y) - \eta S_t\frac{\partial \phi}{\partial S}\Big]d\nu(\eta).\end{aligned}$$

To sum it up, our optimal trading strategy is showed in the following cases

(i) buying stock at the maximum possible rate $l_t = \xi$ and $m_t = 0$ when

$$\frac{\partial V}{\partial x} - aS_t\frac{\partial V}{\partial y} \geq 0 \qquad \text{and} \qquad \frac{\partial V}{\partial x} - bS_t\frac{\partial V}{\partial y} > 0; \tag{2.16}$$

(ii) selling stock at the maximum possible rate $m_t = \xi$ and $l_t = 0$ when

$$\frac{\partial V}{\partial x} - aS_t\frac{\partial V}{\partial y} < 0 \qquad \text{and} \qquad \frac{\partial V}{\partial x} - bS_t\frac{\partial V}{\partial y} \leq 0; \tag{2.17}$$

(iii) doing nothing, that is $m_t = l_t = 0$ when

$$\frac{\partial V}{\partial x} - aS_t\frac{\partial V}{\partial y} \leq 0 \qquad \text{and} \qquad \frac{\partial V}{\partial x} - bS_t\frac{\partial V}{\partial y} \geq 0. \tag{2.18}$$

The above argument shows that the optimization problem (2.13) is a free boundary problem. Besides, the state space $[0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ is partitioned into *buy*, *sell* and *no-transaction* regions, which are characterized by inequalities (2.16), (2.17), and (2.18), respectively. For sufficiently large ξ, the state space remains divided into a *buy region* $\mathcal{B}$, a *sell region* $\mathcal{S}$, and a *no-transaction region* $\mathcal{N}$, and the optimal trading strategy requires an immediate transaction to the boundary of the buy region, $\partial\mathcal{B}$ or that of the sell region $\partial\mathcal{S}$, if the state is in region $\mathcal{B}$ or $\mathcal{S}$. Hence the value function $V^i(t,S,x,y)$ satisfy

$$\begin{aligned} V^i(t,S,x,y) &= V^i(t,S,x+\epsilon_x, y - aS\epsilon_x) \qquad \text{in } \mathcal{B}, \\ V^i(t,S,x,y) &= V^i(t,S,x-\epsilon_x, y + bS\epsilon_x) \qquad \text{in } \mathcal{S}, \end{aligned}$$

in which ϵ_x (the number of shares bought or sold by the investor) can take any positive value up to the number required to take the state to $\partial\mathcal{B}$ or $\partial\mathcal{S}$. Instantaneous transaction from the interior of the buy (or sell) region to the buy (or sell) boundary takes place by letting $\xi \to \infty$. In the no-transaction region $\mathcal{N}$, the value function follows equation (2.15) for the trading strategies $l_t = m_t = 0$ and satisfy inequalities (2.18). Therefore, the above discussion yields the following free boundary problem for the singular stochastic control value function V^i:

$$\begin{cases} \mathcal{L}_1 V^i = 0 & \text{in } \mathcal{N} \\ \dfrac{\partial V^i}{\partial x} - aS\dfrac{\partial V^i}{\partial y} = 0 & \text{in } \mathcal{B} \\ \dfrac{\partial V^i}{\partial x} - bS\dfrac{\partial V^i}{\partial y} = 0 & \text{in } \mathcal{S} \\ V^i(T,S,x,y) = U(y + Z^i(S,x)). & \end{cases} \tag{2.19}$$

2.3 Solutions for the Negative Exponential Utility Function

We further assume that the investor has the negative exponential utility function

$$U(z) = 1 - e^{-\gamma z}, \tag{2.20}$$

in which γ is the constant absolute risk aversion (CARA) parameter. For the equations (2.19), this utility function can reduce much of computational effort and is simple to interpret. Furthermore, the option price based on the exponential utility function is a good approximation to the price implied by any hyperbolic absolute risk aversion utility function with the same level of absolute risk aversion; see Andersen and Damgaard (1999). Davis et al. (1993) show that for the utility function (2.20), the definition of the value function (2.13) can be written as

$$V^i(t, S, x, y) = 1 - \exp\{-\gamma y e^{r(T-t)}\} H^i(t, S, x), \tag{2.21}$$

where

$$\begin{aligned} H^i(t, S, x) &:= \inf_{L,M} E\big\{ \exp[-\gamma(y_T + Z^i(S_T, x_T) - y_t e^r(T-t))] \big| S_t = S, x_t = x \big\} \\ &= 1 - V^i(t, S, x, 0). \end{aligned}$$

Plugging (2.21) into (2.19) and defining the operator $\mathcal{L}_2$ as

$$\mathcal{L}_2\phi := \frac{\partial \phi}{\partial t} + \alpha S \frac{\partial \phi}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 \phi}{\partial S^2} + \int_{-1}^{\infty} \Big[\phi(t, S(1+\eta), x) - \phi(t, S, x) - \eta S \frac{\partial \phi}{\partial S}\Big] d\nu(\eta),$$

we obtain the following free boundary problem for $H^i(t,S,x)$,

$$
\left\{
\begin{array}{rcll}
\mathcal{L}_2 H^i(t,S,x) & = & 0 & x \in [X_b(t,S), X_s(t,S)], \\
\dfrac{\partial H^i}{\partial x}(t,S,x) & = & -a\gamma S e^{r(T-t)} H^i(t,S,x) & x \le X_b(t,S), \\
\dfrac{\partial H^i}{\partial x}(t,S,x) & = & -b\gamma S e^{r(T-t)} H^i(t,S,x) & x \ge X_s(t,S), \\
H^i(T,S,x) & = & \exp\{-\gamma Z^i(S,x)\}, &
\end{array}
\right.
\tag{2.22}
$$

in which $X_b(t,S)$ and $X_s(t,S)$ are the buy and sell boundaries, respectively; see Davis et al. (1993). Moreover, combining (2.21) with (2.14) yields the reservation buying and selling prices

$$
P^b(S,x) \;=\; -\gamma^{-1}\, e^{-rT} \log\left[H^b(0,S,x)/H^0(0,S,x)\right], \tag{2.23}
$$

$$
P^s(S,x) \;=\; \gamma^{-1}\, e^{-rT} \log\left[H^s(0,S,x)/H^0(0,S,x)\right]. \tag{2.24}
$$

In the limit as $\gamma \to \infty$, Bouchard et al. (2001) have shown that the reservation price converges to the sum of the liquidation value of the initial endowment and the super-replication price of the option. The proof of existence and uniqueness of the solutions of the PIDE is very similar with section 5 in Davis et al. (1993). and Lai and Lim (2002).

2.4 Cost-Constrained Minimization of Pathwise Risk

Grinold and Kahn (2000) pointed that "risk is an abstract concept" and other risk measures that depend on the specific investment applications may be more appropriate than variance. Based on this, Lai and Wong (2004) present a new risk measure called "Pathwise Risk" Similarly, we can derive that, under jump-diffusion process framework, the pathwise risk can be expressed as following:

$$R(L,M;t,T)=\sigma^2\int_t^T S_u^2[x_u-\Delta(u,S_u)]^2\mathrm{d}u+\sum_t^T\theta\lambda[xS(V-1)-P(SV,t)+P(S,t)]^2$$

Based on the hedging cost we defined in (2.12), the value functions $V^i(t,S,x)(i=s,b)$ in formula (23) in Lai and Lim (2009) will have the following form under our model.

$$\begin{aligned} V(t,S,x) &= \inf_{L,M}\mathbb{E}\{\int_t^T F(u,S_u,x_u)\mathrm{d}u+\sum_t^T C(u,S_u,x_u)\Delta u+\Psi(L,M;t,T) \\ &\quad +g(S_T,x_T)|S_t=S,x_t=x\} \end{aligned} \tag{2.25}$$

with $F(u,S,x)=\theta\sigma^2S^2[x-\Delta(u,S)]^2$ and $C(u,S,x)=\theta\lambda[xS(V-1)-P(SV,t)+P(S,t)]^2$ for a short call $(i=s)$ or $\theta\sigma^2S^2[x+\Delta(u,S)]^2$ and $C(u,S,x)=\theta\lambda[xS(V-1)+P(SV,t)-P(S,t)]^2$ for a long call$(i=b)$, and $g(S,x)=-Z^s(S,x)$ or $-Z^b(S,x)$, respectively. And the same as that in utility maximization model, the value function of the cost-constrained minimization problem, both for short and long position, is in the state space $[0,T]\times(0,\infty)\times\mathbb{R}$ which is partitioned into (i) a "buy stock" region $\mathbb{B}$, (ii) a "sell stock" region $\mathbb{S}$ and (iii) a "no-transaction" region $\mathbb{N}$, respectively.

Chapter 3

Computational Algorithm

3.1 Lai and Lim's (2009) Algorithm for Diffusion Case

The singular stochastic control problem, whose no-action region is the no-transaction band, is equivalent to an optimal stopping problem. This relationship was the first pointed out by Bather and Chernoff (1966). After that, Karatzas (1983) and Karatzas and Shreve (1984,1985)gave special cases of the connection between singular control and optimal stopping in an analytical way. Peskir and Shiryaev (2006) did a lot of basic research on optimal stopping theory and free-boundary problems, and gave the proofs that the former one can be transferred to the later one under some conditions. And Peskir and Shiryaev (2006) also discussed some mathematical finance problem that can be reformulated as problems of optimal stopping of stochastic processed and solved by reduction to free-boundary problems. As we discussed in Section (1.4), Davis et al. (1993) designed a numerical method based on Markov chain approximation and solved the singular stochastic control problem via solving an appropriate free boundary problem using discrete time dynamic programming algorithm. However, it's computational expensive since it requires the determination of both when to apply control and how much control to apply. Boetius (2005) showed that the sequential stopping problem, which is called as an entry-exit problem, can be transferred to free-boundary problem via Dynkin Game. Based on this connection, Lai and Lim (2009) designed a new algorithm which is called *Coupled Backward Induction Algorithm* to reduce the computational complexities. After using utility maximization approach and transforming it from stochastic control to free-boundary problem via Dynkin Game, the algorithm computes the option price and the optimal hedging. We will introduce the *Coupled Backward Induction Algorithm* in the section because our algorithm is based on it and extended to jump-diffusion case.

With the diffusion only case, the free boundary problem in 2.22 changes to:

$$\begin{cases} \dfrac{\partial H^i}{\partial t} + \alpha S\dfrac{\partial H^i}{\partial S} + \dfrac{\sigma^2 S^2}{2}\dfrac{\partial^2 H^i}{\partial S^2} = 0, & x \in [X_b(t,S), X_s(t,S)], \\ \dfrac{\partial H^i}{\partial x}(t,S,x) = -a\gamma S e^{r(T-t)} H^i(t,S,x), & x \le X_b(t,S), \\ \dfrac{\partial H^i}{\partial x}(t,S,x) = -b\gamma S e^{r(T-t)} H^i(t,S,x), & x \ge X_s(t,S), \\ H^i(T,S,x) = \exp\{-\gamma Z^i(S,x)\}. & \end{cases}$$

Apply the change of vairiable:

$$\rho = \frac{r}{\sigma^2}, \quad s = \sigma^2(t-T), \quad z = \log(S/K) - (\theta - \frac{1}{2})s, \quad \theta = \frac{\alpha}{\sigma^2} \tag{3.1}$$

And then define $h^i(s,z,x) = H^i(t,S,x)$, this free boundary problem can be rewritten as:

$$\begin{cases} \dfrac{\partial h^i}{\partial s} + \dfrac{1}{2}\dfrac{\partial^2 h^i}{\partial z^2} = 0 & x \in [X_b(s,z), X_s(s,z)] \\ \dfrac{\partial h^i}{\partial x}(s,z,x) = -a\gamma K\, e^{z+(\theta-\rho-0.5)s} h^i(s,z,x) & x \leqslant X_b(s,z) \\ \dfrac{\partial h^i}{\partial x}(s,z,x) = -b\gamma K\, e^{z(\theta-\rho-0.5)s} h^i(s,z,x) & x \geqslant X_s(s,z) \\ h^i(0,z,x) = \exp\{-\gamma K A^i(z,x)\}, & \end{cases} \tag{3.2}$$

Note that $\frac{\partial}{\partial s} + +\frac{1}{2}\frac{\partial^2}{\partial z^2}$ is the infinitesimal generator of space-time Brownian motion which means when (s, Z_s) is in the no-transaction region, the dynamic of $h^i(s, Z_s, x_s)$ is driven by the standard Brownian motion $(Z_s, s \le 0)$. In the buy and sell region, it follows from the second and third equation of 3.2 that

$$h^i(s,z,x) = \exp\{-a\gamma K\, e^{z+(\theta-\rho-0.5)s}[x - X_b(s,z)]\}h^i(s,z,X_b(s,z)), \quad x \le X_b(s,z),$$
$$h^i(s,z,x) = \exp\{-b\gamma K\, e^{z+(\theta-\rho-0.5)s}[x - X_s(s,z)]\}h^i(s,z,X_s(s,z)), \quad x \ge X_s(s,z).$$

And then, take $w^i = \partial h^i/\partial x$, we can get free boundary problem for w^i:

$$\begin{cases} \dfrac{\partial w^i}{\partial s} + \dfrac{1}{2}\dfrac{\partial^2 h^i}{\partial z^2} = 0 & x \in [X_b(s,z), X_s(s,z)] \\ w^i(s,z,x) = -a\gamma K\, e^{z+(\theta-\rho-0.5)s}h^i(s,z,x) & x \le X_b(s,z) \\ w^i(s,z,x) = -b\gamma K\, e^{z+(\theta-\rho-0.5)s}h^i(s,z,x) & x \ge X_s(s,z) \\ w^i(0,z,x) = -\gamma K B^i(z,x)h^i(0,z,x), & \end{cases} \tag{3.3}$$

And they also define the second and third equation above as $w_b(s,z,x)$ and $w_s(s,z,x)$ respectively, and define $h(s,z,x)$ and $\widetilde{h}(s,z,x)$ on the 3-dimension grid $(S_\delta, Z_\delta, X_\epsilon)$. Then, to solve the PDE 3.2 and 3.3, Lai and Lim (2009) considered a *Coupled Backward Induction Algorithm*. We will discuss more detain about it in our computational algorithm. Here, we just provide the algorithm briefly.

Algorithm. For $i = 1, 2, \ldots, N$ and $z \in \mathbf{Z}_\delta$:

(i) Starting at $x_0 \in \mathbf{X}_\epsilon$ with $\widetilde{w}(s_i, z, x_0) < \widetilde{w}_b(s_i, z, x_0)$, search for the first $m \in \{1, 2, \ldots, \}$ (denoted by m^* for which $\widetilde{w}(s_i, z, x_0 + m^*\epsilon) \ge \widetilde{w}_b(s_i, z, x_0 + m\epsilon)$ and set $X_b(s_i, z) = x_i + m^*\epsilon$.

(ii) Using similar steps to find $X_s(s_i, z)$.

(iii) For $x \in \mathbf{X}_\epsilon$ outside the interval compute $h(s_i, z, x)$ and $w(s_i, z, x)$

3.2 Algorithm for Our Model

Using the algorithm in previous section, Lai and Lim (2009) successfully solve the PDE for the free boundary problem and compute the option price and option hedging simultaneously. However, when we consider the jump-diffusion case, we are facing to solve PIDE but PDE.

Similar to the variable change in 3.1, we do the change for the variable in 2.22 as

$$\begin{aligned} &\rho = \frac{r}{\sigma^2}, \quad s = \sigma^2(t-T), \quad z = \log(S/K) - (\theta - \beta - \frac{1}{2})s, \\ &\theta = \frac{\alpha}{\sigma^2}, \quad \beta = \frac{1}{\sigma^2}\int_{-1}^{\infty} \eta \mathrm{d}\nu(\eta), \end{aligned} \tag{3.4}$$

and follow the Lai and Lim (2009) step and let $h^i(s, z, x) = H^i(t, S, x)$. Then equation (2.22) can be rewritten as

$$\begin{cases} \mathcal{L}_3 h^i(s,z,x) &= 0 & x \in [X_b(s,z), X_s(s,z)] \\ \dfrac{\partial h^i}{\partial x}(s,z,x) &= -a\gamma K\, e^{z+(\theta-\rho-\beta-0.5)s} h^i(s,z,x) & x \leqslant X_b(s,z) \\ \dfrac{\partial h^i}{\partial x}(s,z,x) &= -b\gamma K\, e^{z(\theta-\rho-\beta-0.5)s} h^i(s,z,x) & x \geqslant X_s(s,z) \\ h^i(0,z,x) &= \exp\{-\gamma K A^i(z,x)\}, \end{cases} \tag{3.5}$$

where the operator $\mathcal{L}_3$ is defined as

$$\mathcal{L}_3\phi := \frac{\partial \phi}{\partial s} + \frac{1}{2}\frac{\partial^2 \phi}{\partial z^2} + \frac{1}{\sigma^2}\int_{-1}^{\infty} \Big[\phi(s, z + \log(1+\eta), x) - \phi(s,z,x)\Big] d\nu(\eta), \tag{3.6}$$

and, corresponding to definitions (2.7)-(2.9) of terminal wealth, we can derive that

$$\begin{aligned} A^0(z,x) &= x\, e^z (a\, \mathbb{I}_{\{x<0\}} + b\, \mathbb{I}_{\{x\geqslant 0\}}), \\ A^i(z,x) &= A^0(z, x - D^i(z)) + D^i(z), \quad \text{for asset settlement} \quad i = s, b, \\ A^i(z,x) &= A^0(z,x) - (e^z - 1) D^i(z), \quad \text{for cash settlement} \quad i = s, b, \end{aligned}$$

with $D^s(z) = \mathbb{I}_{\{z>0\}}$ and $D^b(z) = -\mathbb{I}_{\{z>0\}}$. Note that the operator $\mathcal{L}_3$ can be viewed as the infinitesimal generator of a jump diffusion process defined through the stochastic differential equation,

$$dz_s = d\widetilde{W}_s + d\widetilde{N}_s,$$

where $\widetilde{W}_s$ is a standard Brownian motion and $\widetilde{N}_s$ is a jump process with jump size scaled by $1/\sigma^2$. This implies that while (s, z) is inside the no-transaction region, the dynamics of $h^i(s, z, x_s)$ is driven by the jump diffusion process $\{z_s, s \leq 0\}$. Furthermore, it follows from (3.5) that, in the buy and sell regions,

$$\begin{aligned} h^i(s,z,x) &= \exp\{-a\gamma K\, e^{z+(\theta-\rho-\beta-0.5)s}[x - X_b(s,z)]\} h^i(s,z,X_b(s,z)), \quad x \leq X_b(s,z), \\ h^i(s,z,x) &= \exp\{-b\gamma K\, e^{z+(\theta-\rho-\beta-0.5)s}[x - X_s(s,z)]\} h^i(s,z,X_s(s,z)), \quad x \geq X_s(s,z). \end{aligned} \tag{3.7}$$

As we discussed before, we can solve the **??** directly which is based on the consistent Markov chain approximation. However, the computational complexities make us choose to follow Lai and Lim's (2009) *Coupled Backward Induction Algorithm*, since this numerical scheme can substantially reduce the work load. Follow the the steps of Lai and Lim (2009), instead of solving (3.5) directly, we consider solving two partial integro-differential equations that are much easier than the original problem (3.5). The first partial integro-differential equation has the same form as (3.5) but the boundaries are assumed to be known. The second partial integro-differential equation is a free boundary problem with differential equations in the buy or sell regions are completely specified. Specifically, the second partial integro-differential equation is constructed as follows. Let $w^i = \partial h^i/\partial x$, then w^i satisfies the free boundary problem:

$$\begin{cases} \mathcal{L}_3 w^i(s,z,x) &=& 0 & x \in [X_b(s,z), X_s(s,z)] \\ w^i(s,z,x) &=& -a\gamma K\, e^{z+(\theta-\rho-\beta-0.5)s} h^i(s,z,x) & x \le X_b(s,z) \\ w^i(s,z,x) &=& -b\gamma K\, e^{z+(\theta-\rho-\beta-0.5)s} h^i(s,z,x) & x \ge X_s(s,z) \\ w^i(0,z,x) &=& -\gamma K B^i(z,x) h^i(0,z,x), & \end{cases} \tag{3.8}$$

where $\mathcal{L}_3$ is defined by (3.6) and $B^i(z,x) = \partial A^i(z,x)/\partial x$ is given by

$$\begin{aligned} B^0(z,x) &= A^0(z,x)/x, \\ B^i(z,x) &= B^0(z, x - D^i(z)), \quad \text{for asset settlement} \quad i = s, b, \\ B^i(z,x) &= B^0(z,x), \quad \text{for cash settlement} \quad i = s, b. \end{aligned}$$

In the sequel, we fix $i = 0, s, b$ and drop the superscript i in w^i and h^i for notational simplicity. The purpose of introducing problem (3.8) is that, when the function $h^i(s,z,x)$ is known, (3.8) becomes an optimal stopping problem associated with a Dynkin game (see Lai and Lim, 2009, Section 3), which can be easily solved by backward induction and random walk approximation (see Chernoff and Petkau, 1986, AitSahalia and Lai, 2007 and Lai et al., 2007).

Now, we introduce our coupled backward induction algorithm that solves (3.5) and (3.8) to obtain the buy and sell boundaries $X_b(s,z)$ and $X_s(s,z)$ and the value function $h(s,z,x)$ simultaneously. Note that the algorithm provided by Lai and Lim (2009) is used to deal with the diffusion only case, or with partial differential equation. However, (3.5) and (3.8) are partial integro-differential equations (PIDE), we need to consider a numerical scheme to calculate the jump part or integral part in the equations. Aitsahlia and Runnemo (2007) talked about the American option pricing under the jump-diffusion model, and in that work, they provided a grid search numerical method based on Bernoulli walk to solve PIDE. We combine this recipe with the couple backward induction algorithm, then we obtain our numerical approach to solve the PIDEs (3.5) and (3.8) and we also finish the computation simultaneously. We present our algorithm below:

Given a small $\delta > 0$, we discretize time and space as follows. Let $s_0 = 0$ and $s_j = s_{j-1} - \delta$ for $j \geq 1$ and set

$$\mathbf{S}_\delta = \{s_n, s_n - \delta, \ldots, \delta, 0\},$$
$$\mathbf{Z}_\delta = \{\sqrt{\delta} j \ : \ j \text{ is an integer } \} = \{0, \pm\sqrt{\delta}, \pm 2\sqrt{\delta}, \ldots\}.$$

Note that the possibility of jump in an interval of length δ can be approximated by $\beta\delta$ and hence the possibility of no jump in the interval is $1 - \beta\delta$. Let S_j^δ be the position of the approximating discrete process after the nth transition. On the grid $\mathbf{Z}_\delta$, a jump occurs at the $(j+1)$ transition if $S_{j+1}^\delta - S_j^\delta = \pm l\sqrt{\delta}$ for $l \geq 2$. We then approximate the jump distribution by partitioning the entire real line into intervals of length $\sqrt{\delta}$, whose probability measure is assigned as the probability mass on the midpoints of the intervals. In particular, let F be the cumulative probability distribution function of the jump size Y, we set

$$dF(l) = \begin{cases} F[(l + \frac{1}{2})\sqrt{\delta}] - F[(l - \frac{1}{2})\sqrt{\delta}], & |l| \geq 2, \\ F[(l + \frac{1}{2})\sqrt{\delta}] - F[(-l - \frac{1}{2})\sqrt{\delta}], & l = 0. \end{cases} \tag{3.9}$$

We then assign the following probabilities

$$P\{S^{\delta}_{j+1} - S^{\delta}_{j} = \pm l\sqrt{\delta}\} = \begin{cases} \frac{1}{2}(1-\beta\delta) & \text{if } l = 1, \\ \beta\delta dF(l) & \text{if } l \neq 1. \end{cases}$$

In practice, we need to truncate the set of l to a finite subset $\{l_{\min}, l_{\min}+1, \ldots, l_{\max}\}$. In such case, we set $dF(l_{\min})$ and $dF(l_{\max})$ as

$$\begin{aligned} \mathrm{d}F(l_{\min}) &= P\{Y \leqslant (l_{\min}+0.5)\sqrt{\delta}\}, \\ \mathrm{d}F(l_{\max}) &= P\{Y > (l_{\max}-0.5)\sqrt{\delta}\}. \end{aligned}$$

Specifically, let $T_{\max}$ denote the largest expiration date of interest and take δ and $\epsilon > 0$ such that $N := \sigma^2 T_{\max}/\delta$ is an integer. The problem (3.8) with known $h(s,z,x)$ can be solved by the following backward induction:

Let $\mathbf{X}_{\epsilon} = \{0, \pm\epsilon, \pm 2\epsilon, \ldots\}$. For $i = 1, 2, \ldots, N$,

$$w(s_i,z,x) = \begin{cases} w_b(s_i,z,x) & \text{if } \widetilde{w}(s_i,z,x) < w_b(s_i,z,x), \\ w_s(s_i,z,x) & \text{if } \widetilde{w}(s_i,z,x) > w_s(s_i,z,x), \\ \tilde{w}(s_i,z,x) & \text{otherwise,} \end{cases}$$

where $x \in \mathbf{X}_{\epsilon}$ and

$$
\begin{aligned}
w_b(s,z,x) &= -a\gamma K e^{z+(\theta-\rho-\beta-1/2)}h(s,z,x), && (3.10)\\
w_s(s,z,x) &= -b\gamma K e^{z+(\theta-\rho-\beta-1/2)}h(s,z,x), && (3.11)\\
\widetilde{w}(s_i,z_j,x_m) &= \frac{1}{2}(1-\beta\delta)\big[w(s_{i-1},z_{j+1},x_m)+w(s_{i-1},z_{j-1},x_m)\big]\\
&\quad +\beta\delta\sum_{l=l_{\min},l\neq\pm 1}^{l_{\max}} \mathrm{d}F(l)\cdot w(s_{i-1},z_{j+l},x_m). && (3.12)
\end{aligned}
$$

The boundaries in (3.8) is determined as follows: $X_b(s_i,z)$ is the largest x for which $\widetilde{w}(s_i,z,x) \leq w_b(s_i,z,x)$ and $X_s(s_i,z)$ is the smallest x for which $\widetilde{w}(s_i,z,x) \geq w_s(s_i,z,x)$.

We then solve the problem (3.5) numerically, provided the boundaries $X_b(s,z)$ and $X_s(s,z)$ are given. In such case, the value function in (3.5) with provided boundaries can also be solved by backward induction. In particular, for $z \in \mathbf{Z}_\delta$, define $h(s_i,z,x)$ by (3.7) (with s replaced by s_i) if $x \in \mathbf{X}_\epsilon$ is outside the interval $[X_b(s_i,z), X_s(s_i,z)]$, and let $h(s_i,z,x) = \widetilde{h}(s_i,z,x)$ with

$$
\begin{aligned}
\widetilde{h}(s_i,z_j,x_q) = \frac{1}{2}(1-\beta\delta)\big[h(s_{i-1},z_{j+1},x_q)+h(s_{i-1},z_{j-1},x_q)\big]\\
+\beta\delta\sum_{l=l_{\min},l\neq\pm 1}^{l_{\max}} \mathrm{d}F(l)\cdot h(s_{i-1},z_{j+l},x_q),
\end{aligned}
\tag{3.13}
$$

if $x \in \mathbf{X}_\epsilon \cap [X_b(s_i,z), X_s(s_i,z)]$. Defining $\widetilde{w}_b$ and $\widetilde{w}_s$ as in (3.10) and (3.11) but with h replaced by $\widetilde{h}$, it suggests our coupled backward induction algorithm described below to solve for $X_b(s_i,z)$ and $X_s(s_i,z)$, as well as to compute values of $h(s_i,z,x)$ for $x \in \mathbf{X}_\epsilon \cap [X_b(s_i,z), X_s(s_i,z)]$.

Algorithm. For $i = 1, 2, \ldots, N$ and $z \in \mathbf{Z}_\delta$:

(i) Starting at $x_0 \in \mathbf{X}_\epsilon$ with $\widetilde{w}(s_i, z, x_0) < \widetilde{w}_b(s_i, z, x_0)$, search for the first $m \in \{1, 2, \ldots, \}$ (denoted by m^* for which $\widetilde{w}(s_i, z, x_0 + m^*\epsilon) \geq \widetilde{w}_b(s_i, z, x_0 + m\epsilon)$ and set $X_b(s_i, z) = x_i + m^*\epsilon$.

(ii) Let $x_m = X_b(s_i, z) + m\epsilon$ for $m \in \{1, 2, \ldots\}$. Compute, and store for use at s_{i+1}, $w(s_i, z, x_m) = \widetilde{w}(s_i, z, x_m)$ as defined by (3.12) and $h(s_i, z, x_m) = \widetilde{h}(s_i, z, x_m)$ by (3.13). Search for the first m (denoted by m^*) for which $\widetilde{w}(s_i, z, x_m) \geq \widetilde{w}_s(s_i, z, x_m)$ and set $X_s(s_i, z) = X_b(s_i, z) + m^*\epsilon$.

(iii) For $x \in \mathbf{X}_\epsilon$ outside the interval $[X_b(s_i, z), X_s(s_i, z)]$, compute $h(s_i, z, x)$ by (3.7) and set $w(s_i, z, x) = w_b(s_i, z, x)$ or $w_s(s_i, z, x)$ as defined by (3.10) and (3.11) according to whether $x \leq X_b(s_i, z)$ or $x \geq X_s(s_i, z)$.

Now it's much clearer to notice the advantage of the coupled backward induction algorithm. Comparing to the discrete-time dynamic programming algorithms of Davis et al. (1993), Clewlow and Hodges (1997), and Zakamouline (2006) that need to perform an additional nonlinear optimization to identify the optimal trade size at each time step, the above algorithm avoids such optimization by solving the coupled problem for (w, h) instead of just for h and hence is much easier to implement. The convergence of the algorithm can be shown by using an argument in Lai and Lim (2009).

3.3 Algorithm for Cost-Constrained Minimization Problem

We present the value function $V(t, S, x)$ for cost-constrained minimization problem in the presence of transaction costs in Section 2.4. Now we discuss the how to solve it in the jump-diffusion framework.

First, we apply the change of variable on 2.25 as this:

$$s = \sigma^2(t - T), \quad z = \log(S/K) - (\rho - \beta k - \frac{1}{2})s, \quad \rho = r/\sigma^2, \quad \beta = \lambda/\sigma^2 \tag{3.14}$$

we find that $v(s, z, x) = V(t, S, x)/K$ is the value function of following singular stochastic control problem for the Brownian motion $\{Z_u, u \leqslant 0\}$:

$$\begin{aligned} v(s, z, x) &= \inf_{L,M} \mathbb{E}\{\int_s^0 f(u, Z_u, x_u)\mathrm{d}u + \sum_s^0 c(u, Z_u, x_u)\Delta u + \psi(L, M; s) \\ &\quad + \widetilde{g}(Z_0, x_0)|Z_s = z, x_s = x\} \end{aligned} \tag{3.15}$$

where $f(s, z, x) = F(t, S, x)/K\sigma^2, \quad c(s, z, x) = C(t, S, x)/K\sigma^2, \quad \psi(L, M; s) = \int_{[s,0)} e^{Z_u + (-\beta k - \frac{1}{2})u}(a\mathrm{d}L_u - b\mathrm{d}M_u)$ and $\widetilde{g}(z, x) = g(Ke^z, x)/K$. As shown in Lai and Lim (2009), the function

$$w(s, z, x) = \frac{\partial v}{\partial x}(s, z, x), \qquad (s, z, x) \in [-\sigma^2 T, 0] \times \mathbb{R} \times \mathbb{R}, \tag{3.16}$$

is the value function of the optimal stopping problem associated with the following Dynkin game:

$$\begin{aligned} w(s, z, x) &= \underline{w}(s, z, x) := \sup_{\tau_L \in \mathscr{F}(s,0)} \inf_{\tau_M \in \mathscr{F}(s,0)} I(\tau_L, \tau_M; s, z, x) \\ &= \overline{w}(s, z, x) := \inf_{\tau_M \in \mathscr{F}(s,0)} \sup_{\tau_L \in \mathscr{F}(s,0)} I(\tau_L, \tau_M; s, z, x) \end{aligned} \tag{3.17}$$

where $\mathscr{F}(a, b)$ denotes the set of sopping times taking values between a and b $(> a)$, $f_x(s, z, x) = \partial f(s, z, x)/\partial x$ $c_x(s, z, x) = \partial c(s, z, x)/\partial x$ and $\widetilde{(g)}_x(z, x) = \partial\widetilde{g}(z, x)/\partial x$ are non-decreasing in

x (see(3.23)and (3.24)), and

$$
\begin{aligned}
&I(\tau_1, \tau_2; s, z, x) \\
&= \mathbb{E}\{\int_s^{\tau_1 \wedge \tau_2} f_x(u, Z_u, x_u)\mathrm{d}u + \sum_s^{\tau_1 \wedge \tau_2} c_x(u, Z_u, x_u)\Delta u + G(\tau_1, \tau_2, Z_{\tau_1}, Z_{\tau_2}) \\
&+\widetilde{g}_x(Z_0, x_0)\mathbb{I}_{\{[\}}1]_{\{\tau_1=\tau_2=0\}}|Z_s = z, x_s = x\},
\end{aligned}
$$

$$
G(s_1, s_2, z_1, z_2) = \begin{cases} -ae^{z_1+(-\beta k-\frac{1}{2})s_1} & if \quad s_1 < s_2 < 0, \\ -be^{z_2+(-\beta k-\frac{1}{2})s_2} & if \quad s_2 < s_1 < 0, \end{cases} \tag{3.18}
$$

In addition to relationship (3.16) between the value functions v and w, the optimal continuation region of the Dynkin game (3.17) coincides with the no-transaction region of the singular stochastic control problem (3.15) in the following sense: If (L^*, M^*) is the optimal control of (3.15) and $\tau_L^* = \inf\{u \in [s,0) : L_u^* > 0\} \quad \tau_M^* = \inf\{u \in [s,0) : M_u^* > 0\} \quad (\inf \varnothing = 0)$ then $w(s,z,x) = I(\tau_L^*, \tau_M^*; s, z, x)$ Moreover, letting

$$
w_b(s,z,x) = -ae^{z+(-\beta k-\frac{1}{2})s} \qquad \text{and} \qquad w_s(s,z,x) = -be^{z+(-\beta k-\frac{1}{2})s} \tag{3.19}
$$

w solves the free boundary problem associated with optimal stopping in the Dynkin game:

$$
\frac{\partial w}{\partial s} + \frac{1}{2}\frac{\partial^2 w}{\partial z^2} + f_x + c_x = 0 \qquad \text{in} \quad \mathbb{N}, \tag{3.20a}
$$

$$
w(s,z,x) = w_b(s,z,x) \qquad \text{in} \quad \mathbb{B}, \tag{3.20b}
$$

$$
w(s,z,x) = w_s(s,z,x) \qquad \text{in} \quad \mathbb{S}, \tag{3.20c}
$$

$$
w(0,z,x) = \widetilde{g}_x(z,x). \tag{3.20d}
$$

Since $w(s,z,x)$ is non-decreasing in x, the region $\mathbb{B}$ in (3.20b) has an upper boundary $X_b(s,z)$ which is the largest x for which $w(s,z,x) = w_b(s,z,x)$; this is the *buy boundary.* Similarly, the region $\mathbb{S}$ in (3.20c) has an lower boundary $X_s(s,z)$ which is the smallest x for which $w(s,z,x) = w_s(s,z,x)$; this is the *sell boundary.* Recall that $a \geqslant 1 \geqslant b > 0$.

In view of the functional central limit theorem, we can approximate standard Brownian motion by a symmetric Bernoulli random walk. Since the horizon for problem (3.17) is always 0, only one numerical program for each set of parameters (β, ρ) need be implemented for all expiration dates T. Also, from (3.14), $s = -\sigma^2 T$ at time $t = 0$. Therefore, letting $T_{\max}$ denotes the largest expiration date of interest and taking small positive δ and ϵ such that $N := \sigma^2 T_{\max}/\delta$ is an integer, we can also use backward induction to solve the optimal stopping problems of the type (3.17). For $i = 1, 2, \ldots, N$.

$$w(s_i,z,x) = \begin{cases} w_b(s_i,z,x) & if \quad \widetilde{w}(s_i,z,x) < w_b(s_i,z,x), \\ w_s(s_i,z,x) & if \quad \widetilde{w}(s_i,z,x) > w_s(s_i,z,x), \\ \widetilde{w}(s_i,z,x) & otherwise, \end{cases} \tag{3.21}$$

where $s_0 = 0, s_i = s_{i-1} - \delta$, $z \in \mathbf{Z}_\delta = \{0, \pm\sqrt{\delta}, \pm 2\sqrt{\delta}, \ldots\}$, $x \in \mathbf{X}_\epsilon = \{0, \pm\epsilon, \pm 2\epsilon, \ldots\}$ and

$$\widetilde{w}(s,z,x) = \delta(f_x(s,z,x) + c_x(s,z,x)) + [w(s+\delta, z+\sqrt{\delta}, x) + w(s+\delta, z-\sqrt{\delta}, x)]/2. \tag{3.22}$$

This suggests the backward induction algorithm described below to solve for $X_b(s_i,z)$ as the largest x for which $\widetilde{w}(s_i,z,x) \leqslant w_b(s_i,z,x)$ and for $X_s(s_i,z)$ as the smallest x for which $\widetilde{w}(s_i,z,x) \geqslant w_b(s_i,z,x)$

Algorithm 1. Backward induction Algorithm 1.

For $i = 1, 2, \ldots, N$ and $z \in \mathbf{Z}_\delta$:

1. Starting at $x_0 \in \mathbf{X}_\epsilon$ with $\widetilde{w}(s_i, z, x) < w_b(s_i, z, x)$, search for the first $j \in \{1, 2, \dots\}$ (denoted by j^*) for which $\widetilde{w}(s_i, z, x_0 + j\epsilon) \geqslant w_b(s_i, z, x_0 + j\epsilon)$ and set $X_b(s_i, z) = x_0 + j^*\epsilon$.

2. For $j \in \{1, 2, \dots\}$ let $x_j = X_b(s_i, z) + j\epsilon$. Compute, and store for use at s_{i+1}, $w(s_i, z, x_j) = \widetilde{w}(s_i, z, x_j)$ as defined by (3.22). Search for the first j (denoted by j^*) for which $\widetilde{w}(s_i, z, x_j) \geqslant w_s(s_i, z, x_j)$ and set $X_s(s_i, z) = X_b(s_i, z) + j^*\epsilon$.

3. For $x \in \mathbf{X}_\epsilon$ outside the interval $[X_b(s_i, z), X_s(s_i, z)]$, set $w(s_i, z, x) = w_b(s_i, z, x)$ or $w_s(s_i, z, x)$ as defined by (3.19) according to whether $x \leqslant X_b(s_i, z)$ or $x \geqslant X_s(s_i, z)$.

For the cost-constrained risk minimization problem, the transformations (3.14) lead to (3.15) with

$$f(s, z, x) = \begin{cases} \theta K e^{2z+(2\beta-1)s}[x - D(s, z)]^2 & \text{for short call} \quad (i = s), \\ \theta K e^{2z+(2\beta-1)s}[x + D(s, z)]^2 & \text{for long call} \quad (i = b), \end{cases}$$

$$c(s, z, x) = \begin{cases} \theta\beta[xKe^{z+(\rho-\beta k-\frac{1}{2})s}(V - 1) - p(s, z + \log V) + p(s, z)]^2/K \\ \qquad\qquad\qquad\qquad \text{for short call} \quad (i = s), \\ \theta\beta[xKe^{z+(\rho-\beta k-\frac{1}{2})s}(V - 1) + p(s, z + \log V) - p(s, z)]^2/K \\ \qquad\qquad\qquad\qquad \text{for short call} \quad (i = b), \end{cases}$$

$$v(0, z, x) = -A^i(z, x), \qquad i = s, b,$$

where $D(s, z) = \Delta(t, S)$, $p(s, z) = P(t, S)$ and, corresponding to definitions (**??**)-(**??**) of terminal settlement value,

$$
\begin{aligned}
A^0(z,x) &= x\,e^z(a\,\mathbb{I}_{\{x<0\}} + b\,\mathbb{I}_{\{x\geqslant 0\}}) \\
A^i(z,x) &= A^0(z,x-D^i(z)) + D^i(z) \quad \text{for asset settlement} \quad i=s,b \\
A^i(z,x) &= A^0(z,x) - (e^z-1)D^i(z) \quad \text{for cash settlement} \quad i=s,b
\end{aligned}
$$

with $D^s(z) = \mathbb{I}_{\{z>0\}}$ (short call) and $D^b(z) = -\mathbb{I}_{\{z>0\}}$ (long call). Thus, Algorithm 1 uses

$$
f_x(s,z,x) = \begin{cases} 2\theta K e^{2z+(2\beta-1)s}[x - D(s,z)] & \text{for short call} \quad (i=s), \\ 2\theta K e^{2z+(2\beta-1)s}[x + D(s,z)] & \text{for long call} \quad (i=b), \end{cases} \tag{3.23}
$$

$$
c_x(s,z,x) = \begin{cases} 2\theta\beta e^{z+(\rho-\beta k-\frac{1}{2})s}(V-1)[xKe^{z+(\rho-\beta k-\frac{1}{2})s}(V-1) - p(s,z+\log V) + p(s,z)] \\ \qquad\qquad\qquad\qquad \text{for short call} \quad (i=s), \\ 2\theta\beta e^{z+(\rho-\beta k-\frac{1}{2})s}(V-1)[xKe^{z+(\rho-\beta k-\frac{1}{2})s}(V-1) + p(s,z+\log V) - p(s,z)] \\ \qquad\qquad\qquad\qquad \text{for short call} \quad (i=s), \end{cases} \tag{3.24}
$$

$$
w(0,z,x) = -B^i(z,x), \qquad i=s,b,
$$

where $B^i(z,x) := \partial A^i(z,x)/\partial x$ is given by $B^0(z,x) = A^0(z,x)/x$ and, for $i=s,b$, $B^i(z,x) = B^0(z,x-D^i(z))$ under asset settlement and $B^i(z,x) = B^0(z,x)$ under cash settlement. Note that θK can be treated as single parameter which controls the width of the no-transaction region.

Note that, from (3.22) and (3.23), the algorithm need to know $D(z,s)$ which is the "theoretical delta" in each step. In Lai's diffusion frame work, they just use Black-Scholes formula with adjusted volatility to get this $D(z,s)$. In our model with jump-diffusion framework, we can calculate this "theoretical delta" by Kou's "double-exponential jump-diffusion" option pricing model.

Chapter 4

Numerical Examples and Results

4.1 Discretization of Kou's Model (2002): Double Exponential Jump Diffusion Model

We now perform a numerical study using the proposed algorithm. For illustration purpose, the price process of the underlying is assumed to follow a double exponential jump-diffusion process which was presented by (Kou, 2002). With mean rate of return $\alpha > 0$ and volatility $\sigma > 0$ as we introduce in 1.3:

$$\mathrm{d}S_t = \alpha S_{t-}\mathrm{d}t + \sigma S_{t-}\mathrm{d}W_t + S_{t-}\mathrm{d}\left(\sum_{i=1}^{N_t}(Q_i - 1)\right), \tag{4.1}$$

where α and σ are the expected return and diffusion volatility of the underlying asset, $\{W_t; t \geqslant 0\}$ is a standard Brownian motion with $W_0 = 0$, $\{N_t; t \geq 0\}$ is a Poisson process with rate λ, Q_i is a sequence of independent and identically distributed positive random variables such that the jump $Y = \log Q$ has asymmetric double exponential distribution with density

$$f_Y(y) = \begin{cases} p\eta_1 \, e^{-\eta_1 y} & \text{if } y \geq 0, \\ (1-p)\eta_2 \, e^{\eta_2 y} & \text{if } y < 0. \end{cases}$$

Here $0 \leq p \leq 1$, and p and $1-p$ represent the probability of positive and negative jumps, respectively. The parameters η_1 and η_2 are assumed to satisfy $\eta_1 > 1$ and $\eta_2 > 0$.

As we discussed in Section (1.3.2), this double exponential jump-diffusion model successfully capture the asymmetric leptokurtic feature of the returns' distribution and the "volatility smile" observed in option prices (see Section 3 and Section 5.3 in Kou, 2002). Moreover, from the empirical tests performed in Ramezani and Zeng (1999), we know that the double exponential jump-diffusion model fits stock data even better than the normal jump-diffusion model. Therefore, we choose this specific jump diffusion model provided by Kou (2002) in our numerical studies.

Note that under this specification, (3.9) becomes

$$\mathrm{d}F(l) = \begin{cases} (1-p)(e^{\eta_2 (l+0.5)\sqrt{\delta}} - e^{\eta_2 (l-0.5)\sqrt{\delta}}) & \text{if } l \leq -2, \\ p(e^{-\eta_1 (l-0.5)\sqrt{\delta}} - e^{-\eta_1 (l+0.5)\sqrt{\delta}}) & \text{if } l \geq 2, \\ (1-p)e^{1.5\eta_2\sqrt{\delta}} - pe^{-1.5\eta_1\sqrt{\delta}} & \text{if } l = 0, \end{cases}$$

and the boundaries of the jump distribution is set as follows:

$$\begin{aligned} \mathrm{d}F(l_{\min}) &= (1-p)\, e^{\eta_2 (l_{\min}+0.5)\sqrt{\delta}}, \\ \mathrm{d}F(l_{\max}) &= p\, e^{-\eta_1 (l_{\max}-0.5)\sqrt{\delta}}. \end{aligned}$$

We then study the reservation price of a short call option with $K = 100$ and $T = 0.5$ under the process (4.1). To illustrate the algorithm, we choose $\alpha = r$, which let us characterize the optimal hedging strategy for the utility-maximization approach by a pair (rather than two pairs) of buy-sell boundaries. We further assume that $\alpha = r = 0$ and $\sigma = 0.3$ for the diffusion part of (4.1), which are same as the parameter configuration used by Clewlow and Hodges (1997). For the discretization of time, space and number of shares of stock, we use $\delta = \epsilon = 10^{-4}$ in the following study.

4.2 Impact of Jump Component

We first investigate the impact of jump components on the price of the call option and buy and sell boundaries. In this experiment, we assume the CARA parameter $\gamma = 1$ and transaction cost $\zeta = \mu = 0.01$. To discuss the impact of different rates and directions of jumps, we let $\eta_1 = \eta_2 = 25$ and consider different values of Poisson rate λ and the probability p of positive jumps.

Figures 4.1 and 4.2 show the optimal buy (lower) and sell (upper) boundaries $X_b(t, S)$ and $X_s(t, S)$ for the short call with $p = 0.5$, $S_0 = 90, 100, 110$, and λ with asset settlement and cash settlement, respectively. For each pair of boundaries, the buy-region is below the buy boundary and the sell-region is above the sell boundary. The region between the two boundaries is the no-transaction. Note that the case of $\lambda = 0$ corresponds to the case of having no jumps (or diffusion only). The no-transaction regions in the diffusion only case (the top left panels in Figures 4.1 and 4.2) are wider than those in the jump-diffusion case (the bottom right panel in Figures 4.1 and 4.2). In fact, when the jump rate λ increases, the no-transaction region become narrower, especially when it is close to the expiration date. This fact shows that large price movement narrows down the no-transaction region. In the

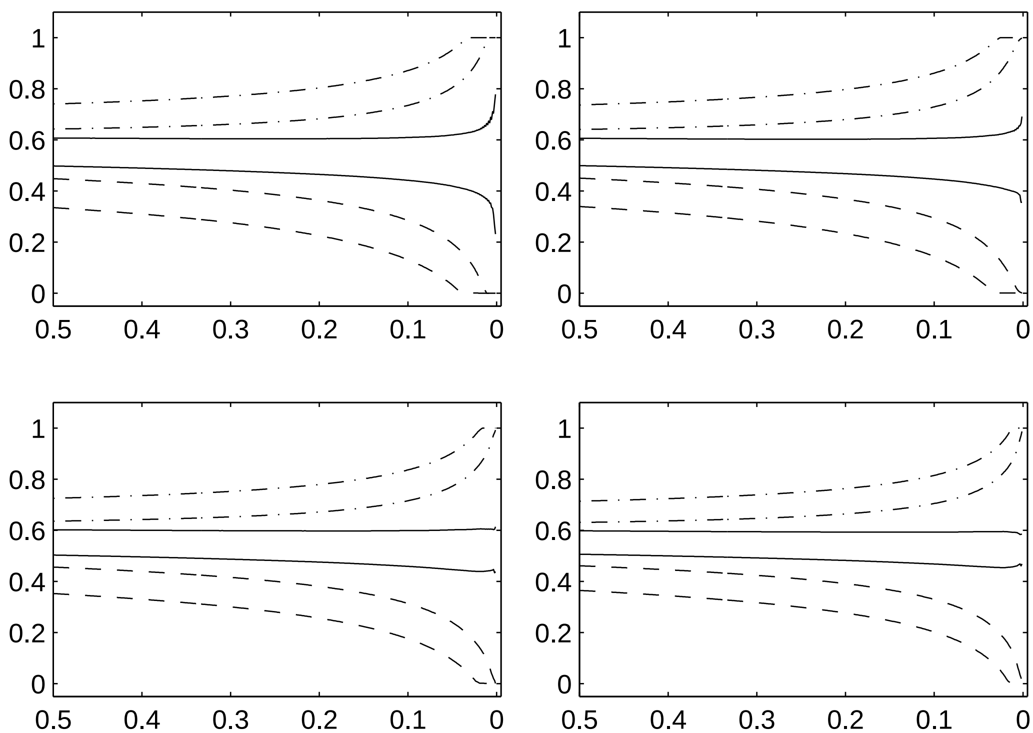

Figure 4.1: Optimal buy (lower) and sell (upper) boundaries $X_b(t, S)$ and $X_s(t, S)$ from CARA utility function for a short call with $\zeta = \mu = 0.01$, $K = 100$, $S = 90, 100, 110$ (dashed, solid and dot-dash lines, respectively) and asset settlement. The Poisson rates in the panels are $\lambda = 0$ (upper left), 1 (upper right), 5 (lower left) and 10 (lower right) and the probability of positive jumps is $p = 0.5$.

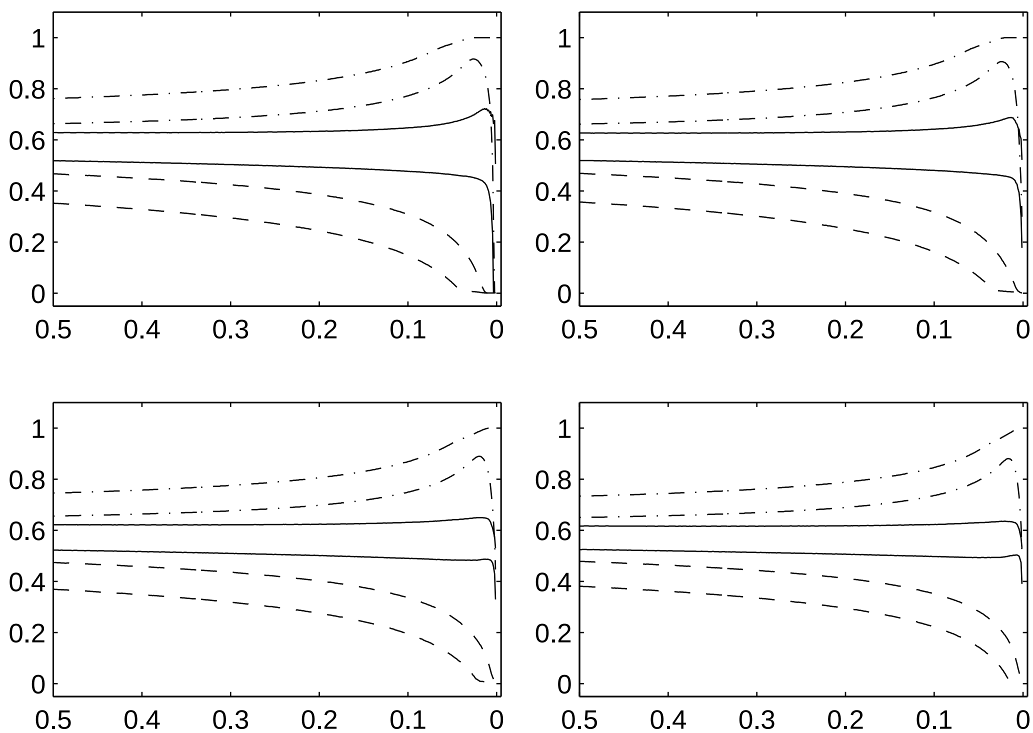

Figure 4.2: Optimal buy (lower) and sell (upper) boundaries $X_b(t, S)$ and $X_s(t, S)$ from CARA utility function for a short call with $\zeta = \mu = 0.01$, $K = 100$, $S = 90, 100, 110$ (dashed, solid and dot-dash lines, respectively) and cash settlement. The Poisson rates in the panels are $\lambda = 0$ (upper left), 1 (upper right), 5 (lower left) and 10 (lower right) and the probability of positive jumps is $p = 0.5$.

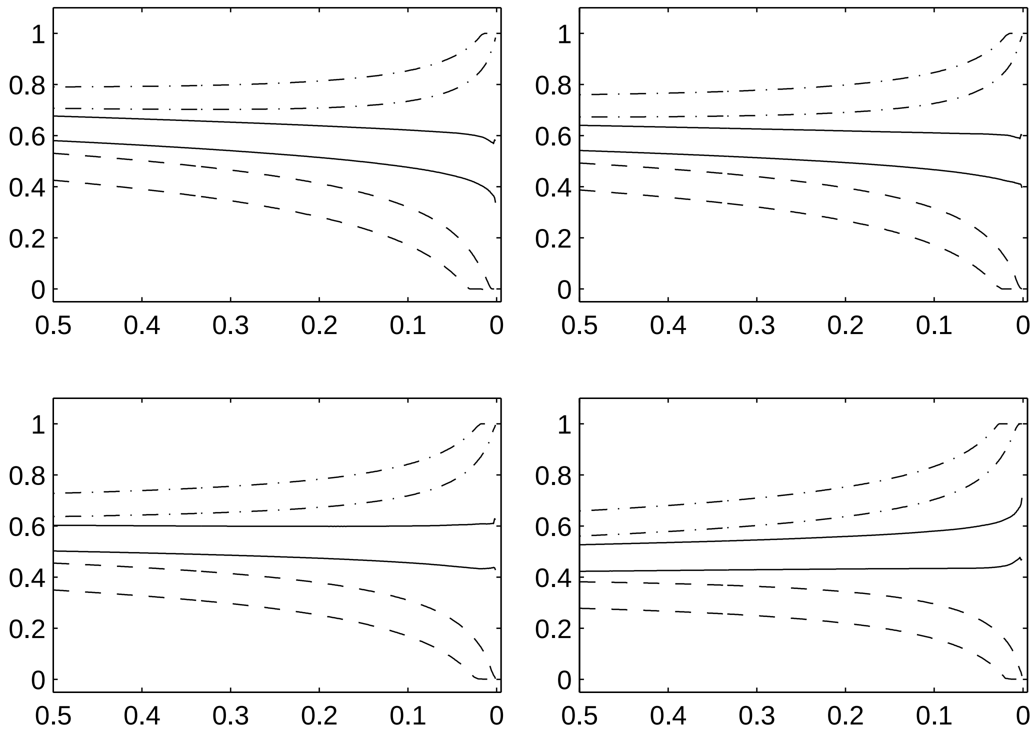

Figure 4.3: Optimal buy (lower) and sell (upper) boundaries $X_b(t, S)$ and $X_s(t, S)$ from CARA utility function for a short call with $\zeta = \mu = 0.01$, $K = 100$, $S = 90, 100, 110$ (dashed, solid and dot-dash lines, respectively) and asset settlement. The probability of positive jumps in the panels are $p = 0.1$ (upper left), 0.3 (upper right), 0.5 (lower left) and 0.9 (lower right) and the rate of jumps is $\lambda = 4$.

asset settlement or Figure 4.1, the boundaries of the out-of-money call option tends to 0, and the boundaries of the in-the-money call option are closer to 1; in the cash settlement or Figure 4.2, the boundaries for call option are tending to 0 before expiration because liquidating any excess position as soon as possible is optimal for the option writer. This phenomenon agrees with the one discussed in Davis et al. (1993).

Figures 4.3 and 4.4 show the optimal buy (lower) and sell (upper) boundaries $X_b(t, S)$ and $X_s(t, S)$ for the short call with $S_0 = 90, 100, 110$, and $p = 0.1, 0.3, 0.5, 0.9$ with asset

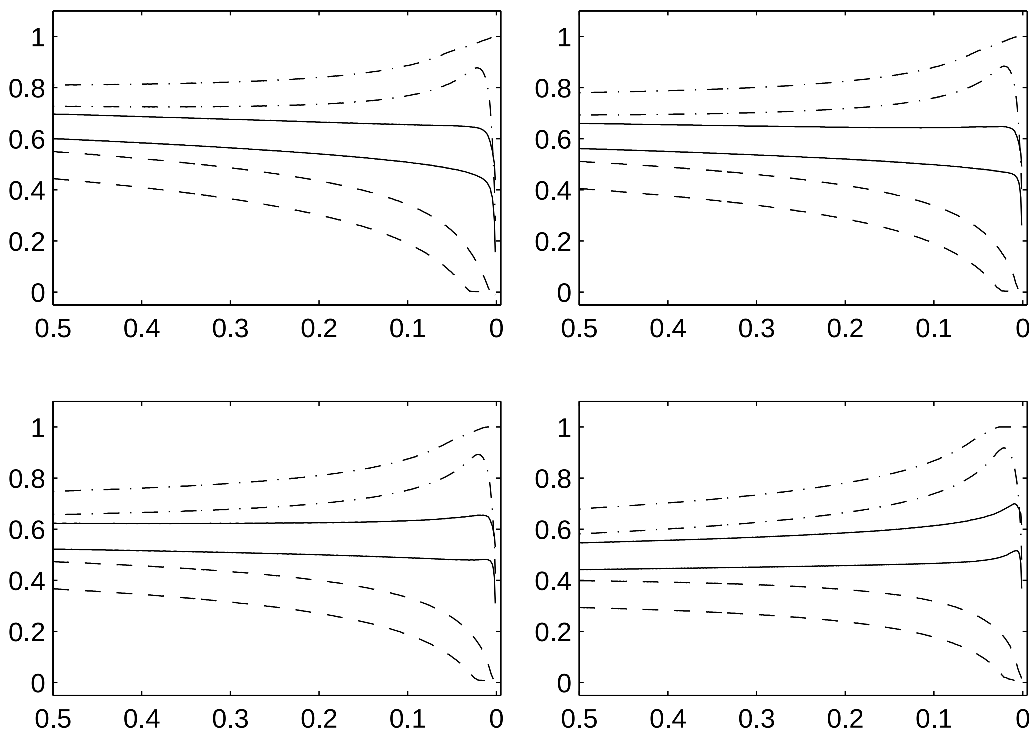

Figure 4.4: Optimal buy (lower) and sell (upper) boundaries $X_b(t,S)$ and $X_s(t,S)$ from CARA utility function for a short call with $\zeta = \mu = 0.01$, $K = 100$, $S = 90, 100, 110$ (dashed, solid and dot-dash lines, respectively) and cash settlement. The probability of positive jumps in the panels are $p = 0.1$ (upper left), 0.3 (upper right), 0.5 (lower left) and 0.9 (lower right) and the rate of jumps is $\lambda = 4$.

Table 4.1: Prices of the short call option under jump-diffusion (4.1).

$p = 0.5$	$\lambda = 0$	$\lambda = 1$	$\lambda = 5$	$\lambda = 10$
Price for Asset Settlement option	5.98	6.13	6.58	7.05
Price for Cash Settlement option	6.31	6.46	6.92	7.39
$\lambda = 4$	$p = 0.1$	$p = 0.3$	$p = 0.5$	$p = 0.9$
Price for Asset Settlement option	9.16	7.75	6.48	4.33
Price for Cash Settlement option	9.59	8.13	6.82	4.59

settlement and cash settlement, respectively. We notice that, as the probability p of positive jumps changes from 0.1 to 0.9, the no-transaction region moves from downward to upward when the time to maturity is closer to 0. Furthermore, when the probability p of positive jumps increases, the stock price S_t has more tendency to move upward, which pushes up the buy and sell boundaries.

Besides the buy and sell boundaries, we also consider the impact of jumps on the prices of the short call options. Table 4.1 shows the prices of the short call option in Figures 4.1-4.4 with $S_0 = 100$. We notice three facts from the table. First, when other conditions are same, the call option price with cash settlement is more expensive than that with asset settlement, this is because the strategy in the cash settlement involves higher hedging cost and hence yields higher option price. Second, more frequent jumps in price (i.e., larger λ) leads to higher option price, as more frequent jumps in price narrow down the no-transaction region and imply higher hedging cost and higher option price. Third, higher probability of positive jumps leads to lower option price, as more frequent positive jumps decreases the number of S_t hitting buying boundaries and hence reduce the hedging cost and option price. To have a better intuition on the impact of jump intensity λ and positive probability p on option price, we show in Figure 4.5 the reservation price for the short call option with $\zeta = \mu = 0.01$, $S = K = 100$, $T = 0.5$ under asset settlement.

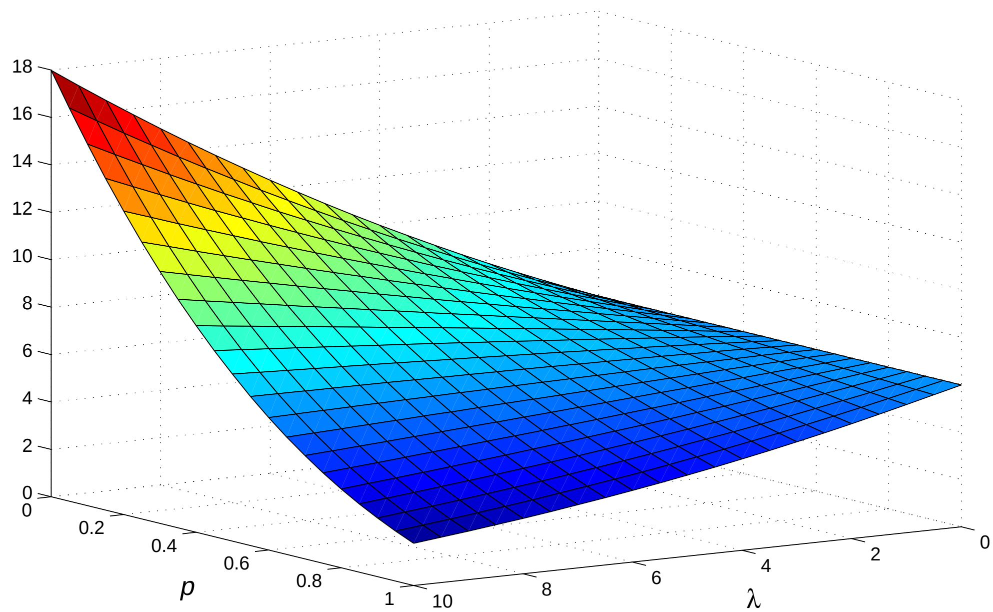

Figure 4.5: Reservation price for short call option with $\zeta = \mu = 0.01$, $S = K = 100$, $T = 0.5$ and asset settlement. The probability of positive $p = 0, \ldots, l$ and the rate of jumps $\lambda = 0, \ldots, 10$.

4.3 Impact of Transaction Cost

We further study the impact of transaction cost on the buy and sell boundaries by considering the following parameter configuration:

$$K = 100, \quad S = 100, \quad \gamma = 1, \quad \lambda = 5, \quad p = 0.5, \quad \eta_1 = 25, \quad \eta_2 = 25$$

in asset and cash settlements, and compute the optimal buy and sell boundaries with the transaction cost $\zeta = \mu = 0, 0.01, \ldots, 0.1$. Figure 4.6 shows the optimal buy and sell boundaries for a short call option with asset and cash settlements. We find that the impact of transaction cost is always monotone for the sell boundary in both asset and cash settlements, while such impact for the buy boundary is monotone in cash settlement but not in asset settlement. Furthermore, when transaction costs converge to 0, the buy and sell boundaries converge to the curve of hedging delta.

4.4 Simulation on Hedging Cost

We now study the total hedging cost of the optimal trading strategies based on the buy and sell boundaries in Section 2. In particular, we consider a short call option with strike $K = 100$ and time to maturity $T = 0.5$. We also assume that the true stock price process is given by the double exponential jump diffusion processes (4.1) with parameters $S_0 = 100$, $\eta_1 = \eta_2 = 25$, and $\lambda = 1, 3, 5, 10, 12$, respectively. The transaction costs are specified by $\zeta = \mu = 0.01$. We then consider trading boundaries computed from a "correctly-" and an "incorrectly-" specified pricing models. The "correctly" specified model uses the jump

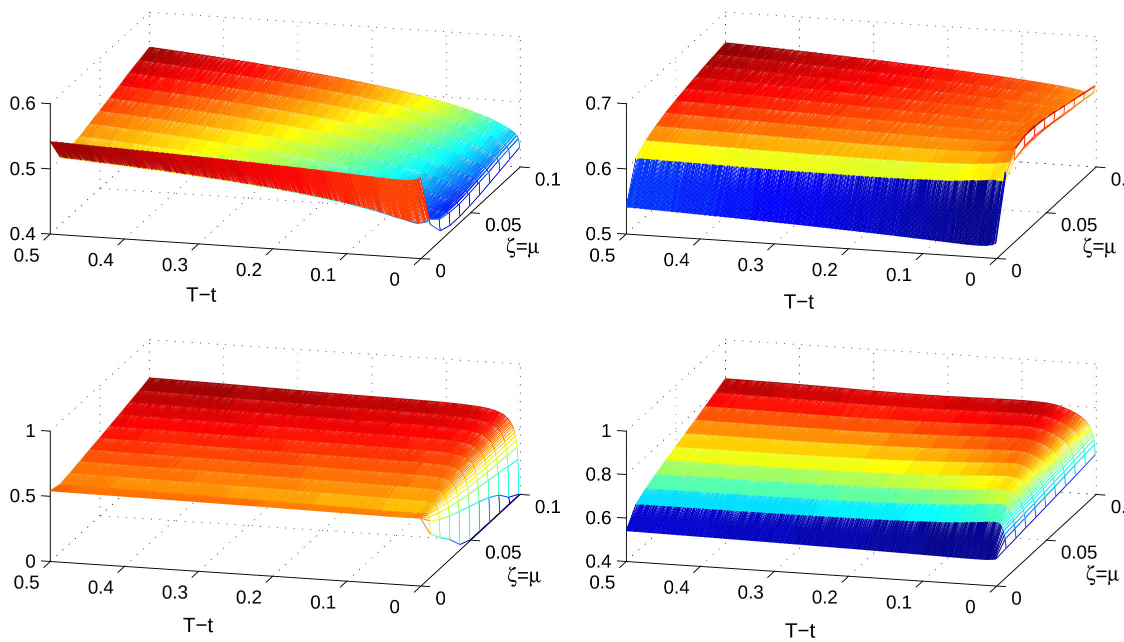

Figure 4.6: Optimal buy (left panels) and sell (right panels) boundaries for a short call with asset (upper panels) and cash (lower panels) settlement. The transaction cost is $\zeta = \mu = 0, 0.01, \ldots, 0.1$.

Table 4.2: Summary statistics of simulation on total hedging cost with $p = 0.5$.

	Correctly-specified model					Incorrectly-specified model				
λ	1	3	5	10	12	1	3	5	10	12
(a) Asset settlement										
$\widehat{C}$	10.255	10.551	10.838	11.517	11.774	10.262	10.558	10.851	11.544	11.801
	(0.004)	(0.005)	(0.005)	(0.007)	(0.007)	(0.004)	(0.005)	(0.005)	(0.006)	(0.007)
$\widehat{\kappa}$	88.24	89.81	91.12	93.70	94.75	85.90	85.95	86.13	86.52	86.57
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
(b) Cash settlement										
$\widehat{C}$	10.734	11.026	11.306	12.000	12.257	10.742	11.049	11.329	12.030	12.302
	(0.004)	(0.005)	(0.006)	(0.007)	(0.007)	(0.004)	(0.005)	(0.005)	(0.006)	(0.007)
$\widehat{\kappa}$	86.13	87.72	89.16	91.85	92.83	84.00	84.14	84.22	84.51	84.66
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)

diffusion process (4.1) with given parameters and is solved by the procedures in Sections 2 and 3; the "incorrectly" specified model assumes no jump in the price process or the price process follow a geometric Brownian motion, $dS_t = \alpha S_t dt + \sigma S_t dW_t$, and is solved by the procedures in Davis et al. (1993). Then for each trading boundaries generated from both models, we simulate 10^5 price paths according to the double exponential jump diffusion processes (4.1), and compute the total hedging cost $\widehat{C}$ according to (2.12) and the total number of trades $\widehat{\kappa}$. We summarize the statistic on $\widehat{C}$ and $\widehat{\kappa}$ from the "correctly-" and the "incorrectly-" specified models in the second and the third columns of Table 4.2, respectively. Note that the correctly specified models always give larger total number of trades but lower total hedging cost, comparing with those of the incorrectly specified models.

4.5 Simulation on Hedging Error

In the presence of transaction costs, we consider to use the mean and standard deviation of the hedging cost (C^s or C^b). Clewlow and Hodges (1997) and Zakamouline (2006) have introduced a similar measure called 'hedging error' which is a linear transformation of the hedging cost defined by:

$$err = \begin{cases} e^{-rT}C^s - P(0, S_0) & \text{for a short call} \\ e^{-rT}C^b + P(0, S_0) & \text{for a long call} \end{cases}$$

Like Lai and Lim (2009), we combine the mean and the standard deviation of err into a single summary statistics

$$\eta := \sqrt{\mathbb{E}\{err^2\}}$$

The table shows the simulation result of mean and standard deviation of hedging error and η under different γ (risk aversion parameter). We consider hedging a short call option with $S = K = 100$, $r = 0$, $\sigma = 0.3$, $T - t = 0.5$ and $\zeta = \mu = 0.01$. In Lai and Lim's (2009) work, they use Black-Scholes formula to compute the market price of option. However, in our jump-diffusion model, we need to use the analytic solution provided by Kou (2002) to get the "market price" $P(t, S)$. The detail of the analytic solution for the option price can be found in Section (1.3.2)

CARA utility maximization:

γ	0.02	0.06	0.2	0.6
Mean	1.40785	1.35393	1.48412	1.7251
SD	6.44691	3.42748	2.62054	2.36339
η	6.59884	3.68521	3.01162	2.92602

Chapter 5

Concluding Remarks

In this dissertation research work, we consider the problem of European option pricing in the presence of proportional transaction cost when the stock price process of the underlying follows a jump diffusion process. Using an approach that is based on maximization of the expected utility of terminal wealth, we transform the option pricing into stochastic optimal control problems, and argue that the value functions of these problems are the solutions of a free boundary problem which consists of a partial integro-differential equation (PIDE) and different boundary conditions.

Since Markov chain approximation method is too complicated involving the jump-diffusion process, we develop a new coupled backward induction algorithm to solve the singular stochastic control problems associated with utility maximization. The algorithm was originally presented by Lai and Lim (2009) which is for the diffusion process framework. We modify this method to make it capable of solving the PIDE we derive, and solving it efficiently by the coupled structure. With the algorithm, we compute the value function and transaction boundaries of the stochastic control problems simultaneously and hence greatly reduces the computational cost, thereby to obtain the option reservation price and hedging boundary which divides the transactions into buy region, sell region and no-transaction region.

In numerical study, we do intensive Monte Carlo simulation on a double-exponential jump diffusion model which was presented by Kou (2002). From the numerical result, we explore the impact of transaction cost and jump on the European option price and hedging strategy. From the comparing of total hedging cost between the diffusion only and jump diffusion process, we see the evidence of advantage of jump-diffusion model.

Part II

Default Risk with Stochastic Covariates in the Presence of Structural Breaks

Chapter 6

Introduction

Default risk, or credit risk, refers to the risk of the event in which companies or individuals will be unable to make the required payments on their debt obligations. And this type of risk is omnipresent in the portfolio of a typical financial institution. For example, the lending and corporate bond portfolios are obviously affected by default risk. Perhaps less obviously, any credit derivative transactions, including OTC(over-the-counter, i.e. non-exchange-guaranteed), such as a swap may substantially affected by the default of one of the parties through the actual pay-off of the transaction. Recently, the research on default risk is getting more and more attention in risk management of the modern finance industrial and academia, especially after the financial crisis.

6.1 Structure Models for Credit Risk

From early 1970s, academia began to consider default risk by using structural models of default timing in which it assumes that a corporation or a firm defaults when its assets drop to a sufficiently low level relative to its liabilities. And the classic structural models for credit risk include Merton (1974), Fisher et al. (1989), and Leland (1994). The idea of this kind of models is proposed by Merton (1974) in which Merton derived the value of option for a defaultable company. In the classical Black-Scholes-Merton model of company debt and equity value, it is assumed that there is a latent firm asset value V determined by the firm's future cash flows, where V follows geometric Brownian motion. Its value at time t, given by V_t, satisfies

$$\frac{\mathrm{d}V_t}{V_t} = r_t \mathrm{d}t + \sigma_t \mathrm{d}W_t,$$

where r_t and σ_t denote asset return rate and volatility of asset value, respectively. W_t, is a standard Wiener process. In Merton (1974)'s work, r_t and σ_t are constants and deterministic. And, he assumes the firm's capital structure just relate to two things: pure equity (that means preference stocks are not considered) and a single zero-coupon debt maturing at time T, of face value B. The default event only occurs when V_T, the asset value at maturity is less than B. Then we have the following payment equations

$$\begin{cases} \text{Receives of debt holders} = \min(V_T, B) \\ \text{Receives of equity holders} = \max(V_T - B, 0) \end{cases}$$

Therefore, the equity holder can be considered as a buyer of European call option in Merton's (1974) work. By assuming there are no dividends, we can use standard Black-Scholes option pricing formula to get the equity market value and the conditional default probability of a firm.

However, Merton (1974)'s model is based on simplified assumptions, and this disadvantage restricts the empirical value of the model. Thus, its subsequent researches mainly focus on relaxing these assumptions. For example, Geske (1977) extends the original single debt maturity assumption to various debt maturities by using compound option modeling. In Leland and Toft (1996)'s work, firms can continuously issue debts which have infinite time to maturity. Comparing with Merton (1974)'s assumption that the default occurs only at the maturity date, another group of structural models (i.e. Black and Cox, 1976) are often referred as *first-passage-time model*. In this class of models, default event can happen not only at the debt's maturity, but also can be prior to that date, as long as the firm's asset value V_t falls below the "barrier B_t" (default trigger value). Thus, the model not only allows valuation of debt with an infinite maturity, but also, more importantly, allows for the default to arrive during the entire life-time of the reference debt or entity. To relaxing the deterministic property of risk-free rate r_t, Longstaff and Schwartz (1995) treat the short-term risk-free interest rate as a stochastic process which converges to long-term risk-free interest rate and is negatively correlated to asset value process, so that the effect of monetary policy to macro economy are considered.

Additionally, the default barrier B_t is also treated dynamically in various papers. For instance, Hui et al. (2003) propose that default barrier should decrease when time goes, since they observe that there is high default risk at time close to maturity. With the observation that firms tend to issue more debt when their asset value increases, Collin-Dufresne and Goldstein (2001) argue that the default trigger value B_t, which was considered as a fixed face value of debt in Merton's (1974) model, should follow a process converging to a fraction of asset value V_t. Actually, this model implied a widespread strategy that firms tend to maintain a constant leverage ratio. Hui et al. (2006) develop this stationary-leverage-ratio model to "incorporate a time-depending target leverage ratio". In the work, they argue that firm's leverage ratio varies across time, because of the movement of initial short-term ratio to long-term target ratio as described in Collin-Dufresne and Goldstein (2001). The most recent Black-Scholes-Merton structural model is proposed by Shibata and Yamada (2009). In their work, they model bank's recovery process for a firm in danger of bankruptcy. When obligor bankrupts, the bank's choice whether the firm should be run or be liquidated affects the losses of the loan.

The most popular practical structural model for default risk is Moody's KMV which set the default trigger value as the risk factor *distance to default* which is the number of standard deviations of annual asset growth by which the asset level (or expected asset level at a given time horizon) exceeds the firm's liabilities. (See Crosbie and Bohn, 2002; Kealhofer, 2003.). From the KMV model, the probability of finishing below "barrier B"(*distance to default* DD) at time T is given as

$$N\left(-\frac{\log(V_t/DD)+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right),$$

This default covariate is a volatility-corrected leverage measure and is built up with market equity data and accounting data for liabilities. We also use the *distance to default* as one of our risk covariate. The detail of construction method for it will be discussed in Section 9.1.2.

There is another classification of the family of structural models. The models can be divided into exogenous default group and endogenous default group. And the grouping only depends on the definitions of default. All the above mentioned works belong to the former group, in which default is defined as when the asset value fall below a trigger value. While the endogenous default models allow obligors choose the time of default strategically. For example, the latter group of models contains Anderson and Sundaresan's (1996) in which the model allows firm to renegotiate the terms of debt contract. When default trigger value is touched, a firm can either bankrupt or give a new but higher interest rate debt contract to debt holder to make the firm continue to run. Tarashev (2005) present an empirical comparison of these two groups of model.

6.2 Intensity Modeling for Credit Risk

In some extent, all models not structural models belong to reduced-form models which is a different type of credit risk model. And it also includes many different type of forms. Empirical firm default analysis can date back at least to Beaver (1966, 1968), and Altman (1968). They are the first to estimate reduced-form statistical models of the likelihood of default on a firm within one accounting period, by identifying accounting data that have statistical explanatory power in differentiating defaulting firms from non-defaulting firms, and they use linear or binomial (such as *logit* or *probit*) models to regress the defaults. Among the covariates, Altaian's *Z-score* is a measure of leverage, defined as the market value of equity divided by the book value of total debt. Once the coefficients of model are estimated, loan applicants are assigned a Z-score to classify they are good or bad. After that, from early 1980s, the empirical work began to focus on qualitative-response models, such as *logit* model and *probit* model. Among these, Ohlson (1980) used an *O-score* method in his year-ahead default prediction model. Then, the most recent generation of reduced-form model for credit risk is dominated by duration analysis, or survival analysis from the viewpoint of statistics. Early in this literature is the work of Lane et al. (1986) on bank default prediction, using time-independent risk factor as covariates. From 1995 when Jarrow and Turnbull began to consider the intensity model (Cox proportional hazard model.), the Cox type counting process model get more and more attention in academia.

In the option-based models we introduced in last subsection, the default event of firms is triggered when firm's assets, or some function thereof, hit or fall below some barrier, or boundary. In contrast of this class of models, the intensity models, which model factors influencing the default event but typically (but not necessarily) leave aside the question of what exactly triggers the default event. Lando (2009) point that there are two main reasons why intensity models are important in the study of default risk. First, the intensity models clearly seem to be the most elegant way of bridging the gap between the models for pricing default risk and credit scoring models or default prediction models (we will introduce in next subsection). Since we can understand the dynamic evolution of the risk covariates and how they influence default probabilities through the intensity-based models which link hazard regressions with standard pricing machinery. And then we can bring the relevant covariates into the pricing models to get the default prediction models. Second, due to the similar mathematical machinery of intensity models and default-free term-structure modeling, the econometric specifications from term-structure modeling and tricks for pricing derivatives can be transferred to defaultable claims, such as basket default swaps, whose equivalent is not readily found in ordinary term-structure modeling, also turn out to be conveniently handled in this setting.

The intensity models are almost exclusively referred to the "Double Stochastic Poisson Process", "Cox Process" or "Counting Process", and were the first presented by Jarrow and Turnbull (1995). In their work, the default process is modeled as a Poisson process N_t with constant intensity λ, the default time τ is exponentially distributed as consequence. Apparently, in this setting, the assumption that the intensity λ is constant over time and across the loan clusters (e.g. across different credit ratings or industries) is very impractical and the following work mainly focus on the concerns to modify the assumption.

Madan and Unal (1998) assume that the intensity λ_t is a nonnegative measurable function of X_t, the excess return on the issuer's equity. And in the setting, the default time τ is

$$\tau = \inf \left\{ t : \int_0^t \lambda(X_u) \mathrm{d}u \geq E \right\}$$

where E is an exponential random variable. Intuitively, the intensity should change over time and differ across the loans' properties. So, in practice, it is a natural idea to allow default intensities to depend on some observable variables which can affect probability of default.

These variables can include firm's accounting information such as EBIT/Asset, market data such as market equity price, macroeconomic variables such as CPI index, and other quantities such as duration of its loan. Based on this consideration, Carling et al. (2007)'s model uses all these kinds of variables, by assuming a linear relationship is held between the selected variables and the log value of intensities. And they found that accounting variables and macroeconomic variables are most powerful to explain the credit risk.

The other kind of thinking is to modify constant intensity to a stochastic process. Lando (1998) denote the time-dependent stochastic intensity λ_t as

$$\mathrm{d}\lambda_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t$$

This equation shows that λ_t follows Brown motion process. μ and σ are the mean and volatility of the intensity; W_t is a standard Wiener process. Furthermore, extending this models by incorporate the dependence between interest rate r_t (modulated by term-structure model) and default intensity λ_t, the following research work illustrates the different structures of affine model. For example, the correlation is through the noise term,

$$\begin{cases} \mathrm{d}r_t &= \kappa_r(r_t - \theta_r)\mathrm{d}t + \sigma_r \mathrm{d}W_t^1, \\ \mathrm{d}\lambda_t &= \kappa_r(\lambda_t - \theta_\lambda)\mathrm{d}t + \sigma_\lambda(\rho \mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2), \end{cases}$$

or correlation is through affine dependence,

$$\begin{cases} \mathrm{d}r_t &= \kappa_r(r_t - \theta_r)\mathrm{d}t + \sigma_r \mathrm{d}W_t^1, \\ \lambda_t &= \alpha + \beta r_t \end{cases}$$

For more on the structure of affine models, see Dai and Singleton (2000). Duffie and Liu (2001) consider negative correlation in the quadratic setting. For links to Heath-Jarrow-

Morton (HJM) modeling, see Schhönbucher (1998), Duffie and Singleton (1999), and Bielecki and Rutkowski (2002). The other intensity models which involve more technical, such as Kusuoka (1999) who consider the risk premium in intensity models and Duffie and Lando (2009) discuss the role of incomplete information. Lee and Urrutia (1996) used a duration model based on a Weibull distribution of default time and they compare duration model and logit models in forecasting insurer insolvency, finding that, for their data, an intensity model identifies more significant variables than does a logit model. Intensity models based on time-varying covariates include those of McDonald and Van de Gucht (1999), in a model of the timing of high yield bond defaults and call exercises.

From a modeling perspective, maybe the analytical tractability of intensity models comes at too high a price, because they can not provide explicit description of default as the first hitting time of a film's asset value. However, in fact, the intensity models arise naturally in a structural model if we notice that the asset value or the distance to default may not be observed perfectly.

6.3 Correlated Default Models and Unobservable Covariates

There is, as pointed by Lando (2009), some loss of generality using defaults modeled as doubly stochastic Poisson processes or Cox processes, and this loss of generality is mainly apparent when modeling multiple defaults by the same firms or defaults of strongly interlinked companies. Since a simultaneous strong change in a common state variable controlling default intensities of several firms can induce high levels of interdependence even while preserving the structure of Cox processes, modeling dependence among default events is one of the biggest challenges of credit risk models. In financial industry, this problem is also a big concern in the credit risk management, because dependence among the default event affects the distribution of loan portfolio losses and is therefore critical in determining quantiles or other risk measures used for allocating capital for solvency purposes, especially facing the emerged market for asset securitization production such as CDOs. So the default correlation which is often used as a general name for dependence has recently received some attention. Lando (2009) summarizes three mechanism for obtaining the dependence:

- Default probabilities are influenced by common background variables which are observable. As in all factor models, we then need to specify the joint movement of the factors and how default probabilities depend on the factors.

- Default probabilities depend on unobserved background variables, and the occurrence of an event causes an updating of the latent variables, which in turn causes a reassessment of the default probability of the remaining events.

- Direct contagion in which the actual default event causes a direct default of another firm or a deterioration of credit quality as a consequence of the default event.

Vasicek (1991) use the mixed binomial model the modulated the correlation of default in dealing with the portfolio losses. And this statistical tool plays a large role in the Basel II process. In the argument of Vasicek (1991), he assumes there is a collection of n firms. The default indicator of firm i is denoted as X_i and it is equal to 1 if firm i defaults before some given time horizon and 0 otherwise. Assume that $\widetilde{p} \in [0,1]$ is a random variable which is independent of all the X_i. Assume that the random variable $X_1, X_2, \ldots, X_n$ are independent and each have default probability $\widetilde{p}$ and denote the density of $\widetilde{p}$ by f and $\mathbb{E}(\widetilde{p})$ by $\overline{p}$. Then we can derive that

$$\mathbb{E}(X_i) = \overline{p} \quad \text{and} \quad Var(X_i) = \overline{p}(1-\overline{p}),$$

and

$$Cov(X_i, X_j) = \mathbb{E}(\widetilde{p}^2) - \overline{p}^2, \quad i \neq j$$

Now if we define the total number of defaults $D_n = \sum_{i=1}^{n} X_i$ for the n firms, then $\mathbb{E}(D_n) = n\overline{p}$ and

$$Var(D_n) = n\overline{p}(1-\overline{p}) + n(n-1)(\mathbb{E}(\widetilde{p}^2) - \overline{p}^2)$$

Based on this mixed binomial model, Davis and Lo (2001) describe a model which incorporates contagion. In binomial-type model, the background variable $\widetilde{p}$ induces the correlation in the default events and it requires assumption of large fluctuations in $\widetilde{p}$ to obtain significant correlation. However, a more direct way to do that is to have direct contagion. Under this model, the default events are divided to two type: direct defaults and defaults triggered through a contagion event. Contagion means that once firm defaults, it may bring down other firms with it. Denote Y_{ij} is an "infection" variable. $Y_{ij} = 1$ means default of firm i immediately triggers default of firm j. Assume all X_i, Y_{ij} are independent Bernoulli variable with $\mathbb{E}(X_i = 1) = p$ and $\mathbb{E}(Y_{ij} = 1) = q$. Then the default indicator of firm i is given as

$$Z_i = X_i + (1 - X_i)\left(1 - \prod_{j \neq i}(1 - X_j Y_{ji})\right).$$

If we denote total number of defaults $D_n = \sum_{i=1}^{n} Z_i$ for the n firms, then we can derive that

$$\begin{cases} \mathbb{E}(D_n) &= n(1-(1-p)(1-pq)^{n-1}) \\ Var(D_n) &= n(n-1)\beta_n^{pq} - (Var(D_n))^2 \end{cases}$$

where $\beta_n^{pq} = p^2 + 2p(1-p)[1-(1-q)(1-pq)^{n-2}] + (1-p)^2[1-2(1-pq)^{n-2} + (1-2pq+pq^2)^{n-2}]$. Other work which emphasizes mixed binomials and common dependence on factor variables include Frey and McNeil (2003), Gordy (2000), and Wilson (1997a,b).

Another very important tool for dealing with default correlation is Copula function which is a multivariate probability distribution for which the marginal probability of each variable is uniformly distributed. Copulas are popular since they allow us to easily model and estimate the distribution of defaults of a basket of firms by estimating marginals and copula separately. Consider a random vector $(X_1, X_2, \ldots, X_n)$. Suppose all margins are continuous, i.e. the marginal CDFs $F_i(x) = \mathbb{P}[X_i \leq x]$ are continuous functions. We know that $F_i(x)$ follows uniform distribution, that is

$$(U_1, U_2, \ldots, U_n) = (F_1(X_1), F_2(X_2), \ldots, F_n(X_n))$$

has uniform margins. The copula of $(X_1, X_2, \ldots, X_n)$ is defined as the joint cumulative distribution function of $(U_1, U_2, \ldots, U_n)$:

$$C(u_1, u_2, \ldots, u_n) = \mathbb{P}[U_1 \leq u_1, U_2 \leq u_2, \ldots, U_n \leq u_n].$$

Copulas are very popular for simulating correlating default times. Assume that the default time of n firms have marginal distribution function $(F_1, F_2, \ldots, F_n)$, which for simplicity we can assume to have an inverse function. Then the simulation of n correlated default times is done by first simulating an outcome $(u_1, u_2, \ldots, u_n)$ on the n-dimensional unit cube following the distribution specified by the preferred copula function, and let the default time be given as $\tau_1 = F_1^{-1}(u_1), \ldots, \tau_n = F_n^{-1}(u_n)$. Some important research work on Copula includes Li (2000), Schubert and Schönbucher (2001) and Rogge and Schönbucher (2003)

Considering correlation in intensity model is another obvious way of handling correlation among a large number of firms. This type of model impose a factor structure on the default intensities. Note that, the goal of factor model is to reduce the parameter specification so as to specify a marginal intensity for each firm and to quantify the part of the marginal intensity which comes from a set of common factors. For example, Duffie and Gârleanu (2001) consider a factor structure model in which the intensity of an individual firm i is decomposed into the two independent components, one coming from a common factor and one being idiosyncratic or firm-specific:

$$\lambda_i(t) = \nu^c(t) + \nu_i^f(t)$$

Now we only consider affine specifications to facilitate the computations. From the basic intensity modeling in Cox process, we can get the survival probability of an individual firm i is

$$\begin{aligned}\mathbb{P}(\tau^i > T) &= \mathbb{E}(\exp\left(\int_0^T \lambda_i(s)\mathrm{d}s\right)) \\ &= \mathbb{E}(\exp\left(\int_0^T \nu_i^f(s)\mathrm{d}s\right))\mathbb{E}(\exp\left(\int_0^T \nu^c(s)\mathrm{d}s\right))\end{aligned}$$

The correlation between firms arises from the common intensity component ν^c. In the work of Lancaster (1990) and Kalbfleisch and Prentice (2002), they treat the common factor as “external” and firm-specific factor as “internal”, that is, cease to be generated once a firm has failed.

So far, all the factors no matter common or firm-specific in the model we discussed are observable covariates. However, the topic of hidden sources of default correlation are getting more attention recently. Collin-Dufresne et al. (2010) and Zhang and Jorion(2007) find that a major credit event at one firm associating with the credit spreads of other firms is consistent with existence of a frailty effect for actual or risk-neutral default probabilities. Collin-Dufresne et al. (2004), Giesecke (2004), and Schönbucher (2003) explore learning from default interpretations, based on the expected effect of unobservable covariates. Das et al. (2007), finds empirical evidence defaults are significantly more correlated than would be suggested by the assumption that default risk is captured by the observable covariates. Duffie et al. (2007) offers an econometric method for estimating term structures of corporate default probabilities over multiple future periods, conditional on firm-specific and macroeconomic covariates. In their model which is without frailty, the observable covariates give substantially better out-of-sample default prediction than does prediction based on credit ratings. After that, Duffie et al. (2009) incorporate the effect of additional unincluded sources of default correlation and show that they have statistically and economically significant implications for the tails of portfolio default loss distributions. The basic idea of the methodology is an application of Bayes's Rule to update the posterior distribution of unobserved risk factors whenever default arrive with a timing that is more or less clustered than would be expected based on the observable risk factors alone. And therefore, this model provides a more realistic assessment of the risk of default.

6.4 Markov Chain Models and Structure Breaks

Jarrow et al. (1997) consider Markov chain in modeling the default events. They treat default event as an absorbing state and default time as the first time when a Markov Chain hits this absorbing state. Extending this model to credit transition model, Nickell et al. (2000) and Feng et al. (2008) fit an ordered probit model and the rating transition probabilities are viewed as functions of latent variables. However, the difference between them is that the former work assumes latent variables derived by observable factors such as industry, residence of the obligor and variables related to business cycle, while Feng et al. (2008) introduce unobservable factors and argue that there's better performance when using unobservable factors. In addition, the ordered probit model can also be applied in sovereign credit migration estimating, as Fuertes and Kalotychou (2007) do.

The estimates of credit rating transition matrices published by rating agencies usually use a discrete-time setting. However, Jarrow et al. (1997) argue that Markov Chain processes can be improved if we adopt the continues-time ones. After that, lot of study was involved in the continuous-time framework such as the the contributions made by Fuertes and Kalotychou (2007), Frydman and Schuermann (2008) and Kadam and Lenk (2008). Monteiro et al. (2006) suggest using"finite non-homogeneous continuous-time semi-Markov process" to model time-dependent matrices. They apply a random transformation on time scale to get a semi-Markov process and show that the nonparametric estimators of the hazard rate functions can be used for consistently estimating these time-dependent transition matrices. Lucas and Klaassen (2006) apply both discrete-time and continuous-time Markov chain model in empirical studies. Frydman and Schuermann (2008) apply Markov mixture model to their analysis, extending it from the original Markov chain model to a mixture of two Markov chains, where the mixing is on the speed of movement among credit ratings. They analyze corporate credit rating history from Standard & Poor's spanning 1981-2002.

Hidden Markov Models (HMM) is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved state. Giampieri et al.(2005) model the occurrence of defaults within a bond portfolio as a simple hidden Markov process. The hidden variable represents the risk state. After obtaining estimates for the model parameters they reconstruct the most likely sequence of the risk state. Banachewicz and Lucas (2008) do a further study on this area, and test the sensitivities of the forecasted quantiles if the underlying Hidden Markov Models is mis-specified.

Different from Giampieri et al. (2005), Xing et al. (2012), in their model, assume that the generators of the rating transition matrices are constant between two adjacent structural change in the economy. And they also assume the generators follow a continuous sate and continuous time nonhomogeneous hidden Markov process since the economy may have infinite regimes. The piecewise constant generators have unobserved, unknown time and unknown number structure breaks which are be modeled as a compounded Poisson process. In detail, the times of structural breaks follow a Poisson process with a constant rate η, hence the duration between two adjacent structural breaks follows an exponential distribution with mean $1/\eta$. The generator matrices between tow adjacent structural breaks are constant and the generator matrix at time t is characterized as $\Lambda(t) = Q_{N_\Lambda(t)}$, where

$Q1, Q2, ...$ are independent and identically distributed random generator matrices such that the off-diagonal elements $\lambda^{(i,j)}$ follow independently a Gamma (α_{ij}, β_i) prior distribution with the density function

$$g(\lambda^{(i,j)}) = \frac{\beta_i^{\alpha_{ij}}}{\Gamma(\alpha_{ij})}[\lambda^{(i,j)}]^{\alpha_{ij}-1}\exp(-\lambda^{(i,j)}\beta_i), \qquad (i,j) \in \mathcal{K}$$

in which $\mathcal{K} = (i,j)|i \neq j, 1 \leq i \leq K-1, 1 \leq j \leq K$. These assumptions allow the model to derive the distributions of the time-varying generators of rating migration matrices and the probability of structural breaks at each time period, given firms' transition history. The derived distribution of generator matrices at a given period is a mixture of Gamma distributions, and the weights of mixture components can be computed explicitly using historical observations. As the number of mixture components changes over time, the model is allowed to incorporate various non-Markovian behaviors in empirical studies. From this perspective, this model extends the mixture model of two independent continuous time homogeneous Markov chains in Frydman and Schuermann (2008).

6.5 Outline

This dissertation research is motivated by the model and ideas mentioned above. In our work, we use an advanced Cox type semiparametric model in survival analysis to modulate default intensity with these two default risk covariates *distance to default* and *firm's trailing stock return*. For firm l and at time t, the model is:

$$\mathrm{d}\mu_l(t) = \exp\{\beta(t)^T Z_l(t)\}\mathrm{d}\mu_0(t),$$

Based on this theoretical foundation and industry practice, it seems natural to consider *distance to default* as a default risk factor or covariate. However, a firm's financial health may have multiple influences over time. For example, firm-specific, sector-wide, and macroeconomic state variables may all influence the evolution of corporate earnings and leverage. As pointed out by Duffie and Lando (2001), if the *distance to default* cannot be accurately measured, then a filtering problem arises, and the default intensity depends on the measured *distance to default* and also on other covariates that may reveal additional information about the firm's conditional default probability. Shumway (2001) bring time-dependent covariates into a discrete duration model and this is computationally equivalent to a multi-period logit model with an adjusted standard-error structure. In predicting one-year default, Hillegeist et al. (2004) also use a discrete duration model. Hillegeist et al. (2004) find, by taking *distance to default*, the theoretical probability of default implied by the Black-Scholes-Merton model, at least in this model setting, can not be entirely explained by *distance to default* And this is supported by Bharath and Shumway (2008) and Campbell et al. (2008), who find that in the presence of market leverage and volatility information, among other covariates, distance to default adds relatively little information. Given the consideration on this empirical study, the usual benefits of parsimony, and especially given the need to model the joint time-series behavior of all default risk covariates chosen, the model of default probabilities estimated in this dissertation work adopts a relatively small set of firm-specific and macroeconomic covariates: *distance to default* and *firm's trailing stock return* which is an important auxiliary covariate suggested by Shumway (2001). Although, like Shumway, we also don't have particular structural interpretation for this covariate, the related research shows that the covariate offers significant incremental explanatory power, perhaps as a proxy for some unobserved risk factor that has an influence on default risk beyond that of the firm's measured *distance to default*. The detail of the two covariates construction will be discussed in Section 9.1.2

We consider the frailty in the model level to consider both observed and unobserved default risk covariates. However, different from the previous models, in our model, we consider the time-varying risk factor coefficients $\beta(t)$. We assume the $\beta(t)$ is piecewise constant and have unknown number and unknown time change-point. And we also construct an estimation procedure for the $\beta(t)$ and hyperparameters which is based on Xing et al. (2012)'s work. This dissertation work is organized as follow. An multiplicative intensity model with unknown structural breaks is introduced in Chapter 7: Section 7.1 introduce the Cox type intensity model; Section 7.2 review the estimation method for the constant β in previous models. And then, we construct an estimation procedure for the time-varying $\beta(t)$ in Section 7.3. Section 7.4 and 7.5 will provide an approximation algorithm and hyperparameter estimation in the time-varying $\beta(t)$ estimation procedure, respectively. In Chapter 8, as a comparison, we introduce smoothing time-varying coefficient model and its estimation procedure. We will provide a real data analysis in Chapter 9: Section 9.1 introduce the real data and covariate we use in the model, including the source, cleaning the construction method. In Section 9.2, we will show the numerical result for the time-varying $\beta(t)$ and other related quantity under structure break setting as Chapter 7. To compared with it, we will show the numerical result for the smoothly time-varying $\beta(t)$ as we talked in Chapter 8. Then a conclusion remarks will be presented in Chapter 10.

Chapter 7

An Multiplicative Intensity Model with Unknown Structural Breaks

7.1 Model Specification

Suppose there are n firms in the study. Let $N_l^\star(t)$ be the number of default event that happened on the lth firm ($l = 1, \ldots, n$) over the time window $[0, t]$ and $Z_l(\cdot)$ be a p-dimensional risk covariate process of the lth firm. Due to the fact that the firm is followed for a limited amount of time, we denote C_l as censoring time for the lth firm. And we also assume that the censoring mechanism is independent with the counting process. So we have

$$\mathbb{E}\{\mathrm{d}N_l^\star(t)|Z_l(t), C_l \geq t\} = \mathbb{E}\{\mathrm{d}N_l^\star(t)|Z_l(t)\}, t \geq 0$$

where $\mathrm{d}N_l^\star(t) = N_l^\star\{(t+dt)^-\} - N_l^\star(t^-)$, the increment of $N_l^\star$ over the small interval $[t, t+\mathrm{d}t)$. Define:

$$N_l(t) = N_l^\star(t \wedge C_l) \qquad \textit{and} \qquad Y_l(t) = I(C_l \geq t),$$

where $a \wedge b = \min(a, b)$, and $I(\cdot)$ is the indicator function.

For any firm over the period $[0, C]$, the observed data consist of $\{N_l(\cdot), Y_l(\cdot), Z_l(\cdot)\}$

$(l = 1, \ldots, n)$. And it's reasonable to assume that $\{N_l(\cdot), Y_l(\cdot), Z_l(\cdot)\}, (l = 1, \ldots, n)$ are independent and identically distributed. Then we denote

$$\mathbb{E}\{\mathrm{d}N_l^{\star}(t) | Z_l(t)\} = \mathrm{d}\mu_l(t) \tag{7.1}$$

and consider the marginal regression model:

$$\mathrm{d}\mu_l(t) = \exp\{\beta(t)^T Z_l(t)\}\mathrm{d}\mu_0(t), \tag{7.2}$$

or

$$\mu_l(t) = \int_0^t \exp\{\beta(u)^T Z_l(u)\}\mathrm{d}\mu_0(u) \tag{7.3}$$

where $\mu_0(t)$ is an unknown continuous function and the time-varying coefficients $\beta(t) \in \mathcal{R}^d$ are *piecewise constant* and have unknown time and unknown number of abrupt change-points. And in detail, to develop an estimation procedure which can incorporate the such feature, we assume that the number of the change-points in $\beta(t)$ follows a Poisson process $\{J(t); t \geq 0\}$ with the rate λ and are independent of $Z_l(t)$ and $C_l, (l = 1, \ldots, n)$. From this assumption, we can derive that duration between two adjacent change-points $\beta(t)$ follows an exponential distribution and $1/\lambda$ is its mean. Between two adjacent change-points of $\beta(t)$ are constant. In our model, the coefficients $\beta(t)$ deliver the impact of risk or default factors to the real default event and from the view of modern credit management, the assumption about $\beta(t)$ we proposed above is reasonable. Since the regression coefficients for the default risk factors shouldn't change often and being characterized by a Poisson process is appropriate. And secondly, coefficients between two consecutive structural change is usually stable which mean there is a stationary relationship between the risk factors and default. Furthermore, we assume that, if a change-point occurs at time t, then the regression coefficient $\beta(t)$ will shift to a new level that is independent of its pre-change values. Specifically, we assume that

$\beta(t) = \beta_{J(t)}$, where $\beta_1, \beta_2, \ldots$ are independent and identically distributed (i.i.d.) normal random vector with mean ν and Σ (can to be estimated from history data.). Note that such assumptions for $\beta(t)$ allows us to consider the case of unknown time and unknown number of change-points, analogous to nonparametric Bayes method using Gamma process prior. And we also note that the model (7.2) specifies how the risk factors or covariates affect the instantaneous rate of the counting process in the presence of structural breaks, and it can degenerates to the case in Lin et al. (2000) when β is constant and there is no change-point during the sample period.

In financial markets, it's very likely that some companies are more prone to the risk coefficients' changes or breaks than others and this kind of heterogeneity is usually characterized by the random-effect model, such as

$$\lambda_Z(t|\eta) = \eta \exp\{\beta(t)^T Z(t)\}\lambda_0(t) \tag{7.4}$$

where η is an unobserved unit-mean positive random variable that is independent of Z. The model (7.2) incorporate this concern (7.4) and integrate out this type of heterogeneity through expectations.

7.2 Estimating Equations For Constant β

To estimate the time-varying coefficients $\beta(t)$, we need to consider a degenerate case first in which $\beta(t)$ is constant and doesn't involve the structural breaks. In particular, we consider an estimating equation for the degenerate case that $\beta(t)$ is a constant random vector with prior distribution $N(\mu, \Sigma)$ given the observations in time window $(t_\star, t^\star)$, which means, $\beta(t) \equiv \beta_0$ for $t \in (t_\star, t^\star)$ and $\beta_0 \sim N(\mu, \Sigma)$. Under this setting, the model (7.2) degenerates to

$$\mu_l(t) = \int_{t_\star}^{t} \exp\{\beta_0^T Z_l(u)\} \mathrm{d}\mu_0(u) \tag{7.5}$$

We know that the model (7.5) is similar to the semiparametric regression models discussed by Pepe and Cai (1993), Lawless and Nadeau (1995), Lawless et al. (1997), and Lin et al. (2000). The difference from their models to model (7.5) is that the latter one only consider a specified time period or segmented time window, while the regression models discussed by Pepe and Cai (1993) and other are for the entire time period $(0, T)$. Due to this reason, we expect that the inference procedures and many others properties developed before can be naturally extended here. Specially, we consider the following counting process argument to construct an estimating procedure of β for the time period $t \in (t_\star, t^\star)$. We let

$$S^{(k)}(\beta, t) = n^{-1} \sum_{l=1}^{n} Y_l(t) Z_l(t)^{\otimes k} \exp\{\beta^T Z_l(t)\}, \quad (k = 0, 1, 2)$$

where $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$ and $a^{\otimes 2} = aa^T$. And let $\overline{Z}(\beta, t) = S^{(1)}(\beta, t)/S^{(0)}(\beta, t)$, and $\overline{z}(\beta, t)$ be the limit of $\overline{Z}(\beta, t)$.

We note that, with model (7.5), the partial likelihood score function for β is $U(\beta, \tau)$, where

$$U_{t_\star, t^\star}(\beta, t) = \Sigma^{-1}(\mu - \beta) + \sum_{l=1}^{n} \int_{t_\star}^{t} [Z_l(u) - \overline{Z}(\beta, u)] \mathrm{d}N_l(u). \tag{7.6}$$

Denote the solution to $U(\beta, \tau) = 0$ by $\beta = \widehat{\beta}(t_\star, t^\star)$ and we then estimate β of model (7.5) by $\widehat{\beta}(t_\star, t^\star)$. To establish the asymptotic distribution of $\widehat{\beta}(t_\star, t^\star)$ under model (7.5), we need to derive the corresponding distribution of $U(\beta, \tau)$. We note that

$$U_{t_\star, t^\star}(\beta_0, t) = \Sigma^{-1}(\mu - \beta_0) + \sum_{l=1}^{n} \int_{t_\star}^{t} [Z_l(u) - \overline{Z}(\beta_0, u)] \mathrm{d}M_{t_\star, t^\star; l}(u), \tag{7.7}$$

where

$$M_{t_\star,t^\star;l}(t) = \int_{t_\star}^{t} \mathrm{d}N_l(u) - \int_{t_\star}^{t} Y_l(u)\exp\{\beta_0^T Z_l(t)\}\mathrm{d}\mu_{t_\star,t^\star;0}(u)$$

Note that for counting process specified via model (7.5), $M_{t_\star,t^\star;l}(t)$ are not martingales, so the martingales central limit theorem is not applicable. However, since

$$\mathrm{d}M_{t_\star,t^\star;l}(t) = I(C_l \geq t \geq t_\star)[\mathrm{d}N_l^\star(t) - \exp\{\beta_0^T Z_l(t)\}\mathrm{d}\mu_0(t)],$$

We have $\mathbb{E}[\mathrm{d}M_{t_\star,t^\star;l}(t)|Z_l(t)] = 0$. Then, using modern empirical process theory, we can show that the process $n^{-1/2}U_{t_\star,t^\star}(\beta,t) \quad (t_\star \leq t \leq t^\star)$ converges weakly to a continuous zero-mean Gaussian process with covariance function

$$\Sigma_{t_\star,t^\star}(s,t) = \mathbb{E}[\int_{t_\star}^{s}\{Z(u) - \overline{z}(\beta_0,u)\}\mathrm{d}M_{t_\star,t^\star}(u)\int_{t_\star}^{t}\{Z(v) - \overline{z}(\beta_0,v)\}\mathrm{d}M_{t_\star,t^\star}(v)] \tag{7.8}$$

where $t_\star \leq s, t \leq t^\star$, and the that $n^{-1/2}(\widehat{\beta}_{(t_\star,t^\star)} - \beta_0)$ is asymptotically zero mean normal vector with covariance matrix

$$\Gamma_{t_\star,t^\star} = A_{t_\star,t^\star}^{-1}\Sigma_{t_\star,t^\star}(t_\star,t^\star)A_{t_\star,t^\star}^{-1} \tag{7.9}$$

in which

$$A_{t_\star,t^\star} = \mathbb{E}[\int_{t_\star}^{t^\star}\{Z(t) - \overline{z}(\beta_0,t)\}^{\otimes 2}Y(t)\exp\{\beta_0^T Z(t)\}\mathrm{d}\mu_0(t)]$$

Please notice that the proof of the argument is essentially the same as the one in Lin et al. (2000) who provided for the whole time period instead of the segmented time period $(t_\star, t^\star)$ here.

To build the covariance matrix $\Gamma_{t_\star,t^\star}$ which is the covariance function of $\widehat{\beta}_{(t_\star,t^\star)}$, we need to calculate the baseline intensity function $\mu_{t_\star,t^\star;0}(\cdot)$. Note that the Aalen-Breslow type estimator is natural estimator here:

$$\widehat{\mu}_{t_\star,t^\star;0}(t) = \int_{t_\star}^{t} \frac{\mathrm{d}\bar{N}(u)}{nS^{(0)}(\widehat{\beta}_{(t_\star,t^\star)}, u)} \quad , \qquad t \in (t_\star, t^\star], \tag{7.10}$$

where $\mathrm{d}\bar{N}(u) = \sum_{l=1}^{n} N_l(u)$. Then applying the strong law of large numbers (Pollard, 1990, page 41) for $\bar{N}(t)$ and the strong consistency of $\widehat{\beta}_{(t_\star,t^\star)}$, we can show that $\widehat{\mu}_{t_\star,t^\star;0}(\cdot)$ converge almost surely to $\mu_0(\cdot)$ an that the covariance matrix Γ can be consistently estimated by 7.9 in which

$$\begin{aligned}
\widehat{A}_{t_\star,t^\star} &= -n^{-1}\partial U_{t_\star,t^\star}(\beta, t^\star)/\partial\beta^T \mid_{\beta=\widehat{\beta}_{(t_\star,t^\star)}} \\
&= -\Sigma^{-1} - n^{-1}\sum_{l=1}^{n}\int_{t_\star}^{t^\star}\{Z_l(u) - \overline{Z}(\widehat{\beta}_{(t_\star,t^\star)}, u)\}^{\otimes 2} Y_l(u)\exp\{\widehat{\beta}_{(t_\star,t^\star)}^T Z_l(u)\}\mathrm{d}\widehat{\mu}_{t_\star,t^\star;0}(u) \\
\widehat{\Sigma}_{t_\star,t^\star} &= n^{-1}\sum_{l=1}^{n}\int_{t_\star}^{t^\star}\{Z_l(u) - \overline{Z}(\widehat{\beta}_{(t_\star,t^\star)}, u)\}\mathrm{d}\widehat{M}_{t_\star,t^\star;l}(u)\int_{t_\star}^{t^\star}\{Z_l(v) - \overline{Z}(\widehat{\beta}_{(t_\star,t^\star)}, v)\}\mathrm{d}\widehat{M}_{t_\star,t^\star;l}(v), \\
\widehat{M}_{t_\star,t^\star;l}(t) &= \int_{t_\star}^{t}\mathrm{d}N_l(u) - \int_{t_\star}^{t} Y_l(u)\exp\{\widehat{\beta}_{(t_\star,t^\star)}^T Z_l(t)\}\mathrm{d}\widehat{\mu}_{t_\star,t^\star;0}(u).
\end{aligned}$$

Note that the marginal regression model (7.5), together with $\mathbb{E}\{\mathrm{d}N^\star(t)|\mathcal{F}_{t-}\} = \mathbb{E}\{\mathrm{d}N^\star(t)|Z(t)\}$, where $\mathcal{F}_{t-}$ represents the σ-field generated by $\{N^\star(s), Z(s)|0 \leq s \leq t\}$, provides a counting process model that is essentially the same as the one in Anderson and Gill (1982). And in their work, they also discuss the asymptotic properties of $n^{-1/2}U(\beta_0, \tau)$ and $n^{1/2}(\widehat{\beta}_{(t_\star,t^\star)} - \beta_0)$. In section 7.3, the likelihood function of model (7.5) will be used for the inference procedure, so we approximate it here by the partial likelihood in this Anderson-Gill type model. In particular, let $t_{t_\star,t^\star;(1)} < \cdots < t_{t_\star,t^\star;(w)}$ denote the w unique ordered default event times in $(t_\star, t^\star]$ and let $\mathcal{G}_k$ denote the set of individuals with an default events at $t_{t_\star,t^\star;(k)}$ The number of individuals in $\mathcal{G}_k$ is denoted as N_k. Then the partial likelihood of risk factors coefficients given observations in $(t_\star, t^\star]$ can be expressed as

$$L_{t_\star,t^\star}(\beta) = \prod_{k=1}^{w} \left\{ \frac{\exp(\sum_{l\in\mathcal{G}_k} \beta^T Z_l(t_{t_\star,t^\star;(k)}))}{[\sum_{l=1}^{n} Y_l(t_{t_\star,t^\star;(k)}) \exp(\beta^T Z_l(t_{t_\star,t^\star;(k)}))]^{N_k}} \right\}; \tag{7.11}$$

In section (7.3), we approximate likelihood function $\psi(\cdot)$ by this partial likelihood (7.11) at $\beta = \widehat{\beta}_{t_{m-1},t_h}$, i.e.,

$$\psi_{t_m,t_h} \approx L_{t_{m-1},t_h}(\widehat{\beta}_{t_{m-1},t_h}) \tag{7.12}$$

7.3 An Estimation Procedure for Time-Varying $\beta(t)$

Now, we develop an estimation procedure for the time-varying coefficients $\beta(t)$. In a discrete time sample period: $0 < t_1 < \cdots < t_H = \overline{C}$ where $\overline{C} = \max(C_1, \ldots, C_n)$. For convenience, we scale the all C_l to make $\overline{C} = 1$. And let $t_h = h/H, (h = 1, \ldots, H)$ and $H \sim O(n^\epsilon)$ for some $\epsilon \in (0, 1)$. We assume that the abrupt change only happen at time $t_1, \ldots, t_H$. We define the variable $J_1 = 1$ and $J_h = J(t_h-) - J(t_{h-1}-)$ for $h = 2, \ldots, H$ to denote whether $\beta(t)$ are the same in the period (t_{h-2}, t_{h-1}) and (t_{h-1}, t_h). Then $\{J_h\}$ is a sequence of independent Bernoulli random variables with success probability $p_h = 1 - \exp(-\lambda/H)$.

Define B_{mhk} as the event that the most recent change-points of $\beta(t)$ happen before and after t_h are at t_{m-1} and t_k, respectively, and $\mathcal{F}_{(t_{m-1},t_k)}$ as the information set consisting of events and covariate history in (t_{m-1}, t_k). That is

$$\mathcal{F}_{(t_{m-1},t_k)} = \{N_l^\star(t), Z_l(t); l = 1, \ldots, n, t \in (t_{m-1}, t_k)\}$$

According the definition of B_{mhk}, the regression coefficients in the intensity model of firms' default is constant, we denote it as $\beta_{(t_m,t_k)}$. Let $U(\beta(t_h)|\mathcal{F}_{(0,t_H)})$ be the partial likelihood score function for $\beta(t_h)$ given all the observations during $(0, t_H)$. Considering the multiple change-points during the whole sample period, we note that $U(\beta(t_h)|\mathcal{F}_{(0,t_H)})$ can be expressed as

$$U(\beta(t_h)|\mathcal{F}_{(0,t_H)}) = \sum_{0\leq m\leq h\leq k\leq H} \pi_{mhk} U(\beta_{(t_{m-1},t_k)}|\mathcal{F}_{(t_{m-1},t_k)}, B_{mhk}) \tag{7.13}$$

in which $\pi_{mhk} = \mathbb{P}(B_{mhk}|\mathcal{F}_{(0,t_H)})$ and it represents the probability of B_{mhk} given all the observation. Equation (7.13) decomposed $U(\beta(t_h)|\mathcal{F}_{(0,t_H)})$ into a mixture of localized estimating equations in which no change-points are involved. This shows us the big picture on the issue of constructing an estimate of $\beta(t_h)$ give the all observed information $\mathcal{F}_{(0,t_H)}$.

First, we need to calculate the mixture probability π_{mhk}, and then conditional on $\mathcal{F}_{(t_{m-1},t_k)}$ and B_{mhk}, we solve the localized estimating equation for $\beta_{(t_{m-1},t_k)}$ which satisfies

$$U(\beta_{(t_{m-1},t_k)}|\mathcal{F}_{(t_{m-1},t_k)}, B_{mhk}) = 0 \tag{7.14}$$

We denote $\beta_{(t_{m-1},t_k)}$ which is the solution of (7.14) as $\widehat{\beta}_{(t_{m-1},t_k)}$. From the structure uncovered by equation (7.13), we can construct the following estimate of $\beta(t_h)$ given $\mathcal{F}_{(0,t_H)}$:

$$\widehat{\beta}(t_h) = \sum_{0\leq m\leq h\leq k\leq H} \pi_{mhk}\widehat{\beta}_{(t_{m-1},t_k)} \tag{7.15}$$

Finally, we extend (7.15) to obtain an estimate for $\beta(t)$, i.e., $\widehat{\beta}(t) = \widehat{\beta}(t_h)$ for $t \in (t_{h-1}, t_h]$

Now, we focus on some details of these steps. To compute π_{mhk}, we use the hidden Markov filtering approach to multiple change-points developed by Lai and Xing (2011). In particular, let $R_h = \max\{t_{m-1}|J_m = 1, m \leq h\}$ and it denote the time of the most recent change-point up to time t_{h-1}. Let $\eta_{m,h} = \mathbb{P}(R_h = t_{m-1}|\mathcal{F}_{(0,t_h)})$. Then the distribution of

$\beta(t_h)$ conditional on $\mathcal{F}_{(0,t_h)}$ or the forward filter is expressed as

$$f(\beta(t_h)|\mathcal{F}_{(0,t_h)}) \sim \sum_{m=1}^{l} \eta_{m,h} f(\beta_{(t_{m-1},t_h)}|\mathcal{F}_{(t_{m-1},t_h)}) \tag{7.16}$$

where $f(\beta_{(t_{m-1},t_h)}|\mathcal{F}_{(t_{m-1},t_h)})$ is the posterior distribution of $\beta(t_h)$ given $R_h = t_{m-1}$ and $\mathcal{F}_{(t_{m-1},t_h)}$, and the mixture probability are expressed as $\eta_{m,h} = \eta^\star_{m,h}/\sum_{m=1}^{h}\eta^\star_{m,h}$ in which

$$\eta^\star_{m,h} = \begin{cases} p_h\psi_{t_h,t_h} & m = h, \\ (1-p_h)\eta_{m,h-1}\psi_{t_m,t_h}/\psi_{t_m,t_{h-1}} & m < h. \end{cases} \tag{7.17}$$

Proof. First, we derive the posterior distribution of $\beta(t_h)$ given $\mathcal{F}_{(0,t_h]}$ (7.16). Note that conditional on $J_h = 1$ or 0, we have

$$\begin{aligned} &f(\beta(t_h)|\mathcal{F}_{(0,t_h]}) \propto f(\beta(t_h), \mathcal{F}_{(t_{h-1},t_h]}|\mathcal{F}_{(0,t_{h-1}]}) \\ &= p_h f(\beta(t_h), \mathcal{F}_{(t_{h-1},t_h]}|\mathcal{F}_{(0,t_{h-1}]}, J_h = 1) + (1-p_h) f(\beta(t_h), \mathcal{F}_{(t_{h-1},t_h]}|\mathcal{F}_{(0,t_{h-1}]}, J_h = 0) \end{aligned} \tag{7.18}$$

Note the first term that

$$\begin{aligned} & p_h f(\beta(t_h), \mathcal{F}_{(t_{h-1},t_h]}|\mathcal{F}_{(0,t_{h-1}]}, J_h = 1) \\ = \quad & \eta^\star_{h,h} f(\beta(t_h)|\mathcal{F}_{(0,t_h]}, J_h = 1) \\ = \quad & \eta^\star_{h,h} N(\widehat{\beta}_{t_{h-1},t_h}, \Sigma_{t_{h-1},t_h}) \end{aligned}$$

in which

$$\eta^\star_{h,h} = p_h f(\mathcal{F}_{(t_{h-1},t_h]}|\mathcal{F}_{(0,t_{h-1}]}, J_h = 1) = p_h \int f(\mathcal{F}_{(t_{h-1},t_h]}|\beta(t_h)) f(\beta(t_h)) \mathrm{d}\beta(t_h) = p_h\psi_{t_h,t_h}$$

where the likelihood function $\psi(\cdot)$ is approximated by the partial likelihood function defined in (7.10). The second term in (7.18) can be expanded as

$$
\begin{aligned}
& (1-p_h) f(\beta(t_h), \mathcal{F}_{(t_{h-1},t_h]} | \mathcal{F}_{(0,t_{h-1}]}, J_h = 0) \\
= \; & (1-p_h) \sum_{m=1}^{h-1} \mathbb{P}(R_{h-1} = t_{m-1} | \mathcal{F}_{(0,t_{h-1}]}, J_h = 0) f(\beta(t_h), \mathcal{F}_{(t_{h-1},t_h]} | R_{h-1} = t_{m-1}, \mathcal{F}_{(0,t_{h-1}]}, J_h = 0) \\
= \; & \sum_{m=1}^{h-1} \eta^{\star}_{m,h-1} f(\beta(t_h) | R_{h-1} = t_{m-1}, \mathcal{F}_{(0,t_h]}, J_h = 0) \\
= \; & \sum_{m=1}^{h-1} \eta^{\star}_{m,h-1} N(\widehat{\beta}_{t_{m-1},t_h}, \Sigma_{t_{m-1},t_h})
\end{aligned}
$$

in which

$$
\begin{aligned}
\eta^{\star}_{m,h-1} \; = \; & (1-p_h)\eta_{m,h-1} f(\mathcal{F}_{(t_{h-1},t_h]} | R_{h-1} = t_{m-1}, \mathcal{F}_{(0,t_{h-1}]}, J_h = 0) \\
& = (1-p_h)\eta_{m,h-1} \frac{f(\mathcal{F}_{(t_{m-1},t_h]}, R_h = t_{m-1})}{f(\mathcal{F}_{(t_{m-1},t_{h-1}]}, R_{h-1} = t_{m-1})} \\
& = (1-p_h)\eta_{m,h-1} \psi_{t_m,t_h} / \psi_{t_m,t_{h-1}}
\end{aligned}
$$

Hence combining above equations yields the posterior distribution (7.16) and (7.17). □

From our assumption, we know that the multiple change-point process (restricted on the grid $t_0, \ldots, s_H$) is a hidden Markov chain with a stationary distribution. Then we can reverse time and obtain the corresponding backward filter which is analogous to the forward filter. In detail, let $\widetilde{R}_{h+1} = \min(t_k | J_k = 1, k > h)$ and $\widetilde{\eta}_{k,h+1} = \mathbb{P}(\widetilde{R}_{h+1} = t_k | \mathcal{F}_{t_{h+1},t_H})$. The distribution of $\beta(t_h)$ conditional on $\mathcal{F}_{(t_h,t_H)}$ or the backward filter is given by

$$
f(\beta(t_h) | \mathcal{F}_{(t_h,t_H)}) \sim p_h f(\beta(t_h) | \mathcal{F}_0) + (1-p_h) \sum_{k=h+1}^{H} \widetilde{\eta}_{k,h+1} f(\beta_{(t_h,t_k)} | \mathcal{F}_{(t_h,t_k)}), \tag{7.19}
$$

in which $f(\beta(t_h) | \mathcal{F}_0)$ represents the density of $\beta(t_h)$ without any observations (the prior distribution of post-change value of $\beta(t)$), and the mixture probability $\widetilde{\eta}_{k,h+1} = \widetilde{\eta}^{\star}_{k,h+1} / \sum_{k=h+1}^{H} \widetilde{\eta}^{\star}_{k,h+1}$ and

$$\widetilde{\eta}^{\star}_{k,h+1} = \begin{cases} p_{h+1}\psi_{t_{h+1},t_{h+1}} & k = h+1, \\ (1-p_{h+1})\widetilde{\eta}_{k,h+2}\psi_{t_{h+1},t_k}/\psi_{t_{h+2},t_k} & k > h+1. \end{cases} \tag{7.20}$$

Proof. First, we reverse time and note that $\widetilde{J}_h = J_{H-h+1}$ are still i.i.d. Bernoulli and that the time-reversed Markov chain $\widetilde{\beta}(t_h) = \beta(t_{H-h+1})$ has the same transition probabilities as the Markov chain $\beta(t_h)$. In other words, $\{\beta(t_h)\}$ is a reversible Markov chain. Moreover, its stationary distribution is $N(\mu, \Sigma)$. Using the similar argument as the proof for (7.16) and (7.17), we can prove the posterior distribution of $\beta(t_h)$ given $\mathcal{F}_{(t_h,t_H)}$ is obtained by the (7.19) and (7.20). $\square$

Then we use the Bayes Theorem to combine function (7.16) and (7.19) to get the distribution of $\beta(t_h)$ conditional on $\mathcal{F}_{(0,t_H)}$,

$$f(\beta(t_h)|\mathcal{F}_{(0,t_H)}) \sim \sum_{1\le m\le h\le k\le H} \pi_{mhk} f(\beta_{(t_{m-1},t_k)}|\mathcal{F}_{(t_{m-1},t_k)}), \tag{7.21}$$

and the mixture weight π_{mhk},

$$\pi_{mhk} = \pi^{\star}_{mhk} / \sum_{1\le i\le h\le j\le H} \pi^{\star}_{ihj}, \tag{7.22}$$

where

$$\pi^{\star}_{mhk} = \begin{cases} p_h\eta_{m,h} & m \le h = k, \\ (1-p_h)\eta_{m,h}\widetilde{\eta}_{k,h+1}\psi_{t_m,t_k}/(\psi_{t_m,t_h}\psi_{t_{h+1},t_k}) & m \le h < k. \end{cases} \tag{7.23}$$

Proof. Let $f(\cdot|\mathcal{F}_{(t_h,t_H)})$ and $f(\cdot|\mathcal{F}_{(0,t_H)})$ denote the density function of $\beta(t_h)$ given $\mathcal{F}_{(t_h,t_H)}$ and $\mathcal{F}_{(0,t_H)}$, respectively. And let f denote the stationary density function of $\beta(t_h)$ which is the same as the prior Normal random vector $N(\mu, \Sigma)$.

Note the assumption that outcomes are conditionally independent in the time period $(t_{h-1}, t_h))$ given covariates and $\beta(t_h)$ and we can obtain that

$$
\begin{aligned}
& f(\beta(t_h)|\mathcal{F}_{(0,t_H)}) \propto f(\beta(t_h))f(\mathcal{F}_{(0,t_H)}|\beta(t_h)) \\
\propto \quad & f(\beta(t_h))f(\mathcal{F}_{(0,t_h)}|\beta(t_h))f(\mathcal{F}_{(t_{h-1},t_h)}|\beta(t_h)) \\
\propto \quad & f(\mathcal{F}_{(0,t_h)}, \beta(t_h))f(\mathcal{F}_{(t_{h-1},t_h)}, \beta(t_h))/f(\beta(t_h)) \\
\propto \quad & f(\beta(t_h)|\mathcal{F}_{(0,t_h)})f(\beta(t_h)|\mathcal{F}_{t_h,t_H)})/f(\beta(t_h))
\end{aligned}
$$

in which $f(\beta(t_h)|\mathcal{F}_{(0,t_h)})$ and $f(\beta(t_h)|\mathcal{F}_{(t_h,t_H)})$ are the posterior distributions of $\beta(t_h)$ given $\mathcal{F}_{(0,t_h)}$ and $\mathcal{F}_{(t_h,t_H)}$, respectively. We put (7.16) and (7.19) and Normal prior into the above equation and notice the following fact

$$
\begin{aligned}
& N(\widehat{\beta}_{t_{m-1},t_h}, \Sigma_{t_{m-1},t_h}) \cdot N(\widehat{\beta}_{t_h,t_k}, \Sigma_{t_h,t_k})/N(\mu, \Sigma) \\
= \quad & \frac{\psi_{t_m,t_k}}{\psi_{t_m,t_h}\psi_{t_{h+1},t_k}} N(\widehat{\beta}_{t_{m-1},t_k}, \Sigma_{t_{m-1},t_k})
\end{aligned}
$$

then we can arrive at (7.21) and (7.23). □

Now, combining with the localized estimation procedure we discussed in Section 7.2 for period (t_{m-1}, t_k), we can compute the time-varying risk factors' coefficients $\beta(t)$.

7.4 A Bounded Complexity Mixture Algorithm

In our study, one of the purposes is to estimate the times of change points happened in the coefficients. we prefer a fine grid or a large H in practical analysis. However, this grid structure makes the number of mixtures in (7.15) increase with H, resulting unbounded computational complexity and memory requirements in estimating $\beta(t_h)$ as h changes from 1 to H. To address this issue, we follow Lai & Xing (2011) and we consider a *bounded complexity* *mixture*(BCMIX) approximation algorithm which has linear computational complexity by keeping only a fixed number of mixtures.

Now we discuss this approximation algorithm BCMIX in detail. At each t_h, we keep only the most recent b ($1 < b < B$) weights $\eta_{m,h}$ (with $h - b < m \leq h$) and the largest $B - b$ of the $h - b$ remaining weights in (7.16). Denote $\mathcal{K}_{h-1}$ the set of indices j for which $\eta_{j,h-1}$ in (7.16) is kept at time t_{h-1}; then $\mathcal{K}_{h-1} \in \{h-1, \dots, h-b\}$. At time t_h, calculate $\eta^\star_{m,h}$ by (7.17) for $h \in \{h\} \cup \mathcal{K}_{h-1}$ and let j_h be the index not belonging to $\{h, \dots, h-b+1\}$ such that $\eta^\star_{j_h,h} = \min\{\eta^\star_{j,h} : j \in \mathcal{K}_{h-1} \quad \text{and} \quad j \leq h-b\}$. Choosing j_h to be the minimizer farthest from h if the above set has two or more minimizers. Define $\mathcal{K}_h = \{h\} \cup (\mathcal{K}_{h-1} - \{j_h\})$ calculate $\eta_{m,h}$ by

$$\eta_{m,h} = (\eta^\star_{m,h} / \sum_{i \in \mathcal{K}_h} \eta^\star_{i,h}), \qquad \text{for} \quad m \in \mathcal{K}_h$$

We then get an approximation to (7.16). Similarly, for the backward filter, let $\widetilde{\mathcal{K}}_{h+1}$ denote the set of indices $\widetilde{j}$ for which $\widetilde{\eta}_{\widetilde{j},h+1}$ in (7.19) is kept at time t_{h+1}; then $\widetilde{\mathcal{K}}_{h+1} \in \{h+1, \dots, h+b\}$. At time t_h, calculate $\widetilde{\eta}^\star_{k,h}$ by (7.20) for $h \in \{h\} \cup \widetilde{\mathcal{K}}_{h+1}$ and let $\widetilde{j}_h$ be the index not belonging to $\{h, \dots, h+b-1\}$ such that $\widetilde{\eta}^\star_{\widetilde{j}_h,h} = \min\{\widetilde{\eta}^\star_{\widetilde{j},h} : \widetilde{j} \in \widetilde{\mathcal{K}}_{h+1} \quad \text{and} \quad \widetilde{j} \geq h+b\}$. Choosing $\widetilde{j}_h$ to be the minimizer farthest from h if the above set has two or more minimizers. Define $\widetilde{\mathcal{K}}_h = \{h\} \cup (\widetilde{\mathcal{K}}_{h+1} - \{\widetilde{j}_h\})$ calculate $\widetilde{\eta}_{k,h}$ by

$$\widetilde{\eta}_{k,h} = (\widetilde{\eta}^\star_{k,h} / \sum_{i \in \widetilde{\mathcal{K}}_h} \widetilde{\eta}^\star_{i,h}), \qquad \text{for} \quad k \in \widetilde{\mathcal{K}}_h$$

We then get an approximation to (7.19). The approximation to (7.21) can be obtained by combining the above approximation to (7.16) and (7.19) using the argument in the last section.,

$$f(\beta(t_h)|\mathcal{F}_{(0,t_H)}) \sim \sum_{m \in \mathcal{K}_h, k \in \widetilde{\mathcal{K}}_{h+1}} \pi_{mhk} f(\beta_{(t_{m-1},t_k)}|\mathcal{F}_{(t_{m-1},t_k)}), \tag{7.24}$$

in which $\pi_{mhk} = \pi^\star_{mhk} / \sum_{i\in\mathcal{K}_h, j\in\widetilde{\mathcal{K}}_{h+1}} \pi^\star_{ihj}$ and $\pi^\star_{mhk}$ can be obtained by (7.23) for $m \in \mathcal{K}_h, k \in \widetilde{\mathcal{K}}_{h+1}$. Therefore, the time-varying coefficients defined in (7.15) can be approximated by

$$\widehat{\beta}(t_h) \approx \sum_{m\in\mathcal{K}_h, k\in\widetilde{\mathcal{K}}_{h+1}} \pi_{mhk} \widehat{\beta}_{(t_{m-1}, t_k)} \tag{7.25}$$

Naturally, how to choose B and b is becoming our next concern of using the above algorithm. Since the asymptotic properties holds for any fixed pair $(b, B), (b < B)$. Specifically, in the preparation of numerical study, we have tested and compared the performance of $(b, B) = (10, 20), (15, 25)$ and $(25, 50)$ via simulation, and the result for estimated values don't show much difference. To balance the computation time cost and performance, we decide to use $(b, B) = (15, 25)$ in our numerical study which will be presented in Section 9.2.

7.5 Estimation of Hyperparameters

The estimation procedures in Sections 7.2-7.4 contains hyperparameters $\Phi = \{\lambda, \mu, \Sigma\}$. In practical analysis, Φ is unknown and need to estimated from the real data. From the definition (7.17) of $\eta^\star_{m,h}$, it follows that the conditional likelihood of $\mathcal{F}_{(t_{h-1}, t_h)}$ given $\mathcal{F}_{(0, t_{h-1})}$ is

$$f(\mathcal{F}_{(t_{h-1}, t_h)} | \mathcal{F}_{(0, t_{h-1})}) = \sum_{m=1}^{h} \eta^\star_{m,h}$$

in which $\eta^\star_{m,h}$ are functions of the hyperparameter Φ. Given Φ and the observed data $\mathcal{F}_{(0, t_H)}$, the log-likelihood function is expressed as

$$l(\Phi) = \sum_{h=1}^{H} \log f(\mathcal{F}_{(t_{h-1}, t_h)} | \mathcal{F}_{(0, t_{h-1})}) = \sum_{h=1}^{H} \log\{\sum_{m=1}^{h} \eta^\star_{m,h}\} \tag{7.26}$$

Notice the semiparametric feature in our model, we cannot compute this likelihood exactly. Therefore, we compute $\eta^{\star}_{m,h}$ via approximation of ψ_{t_m,t_h} as stated in the equation (7.12).

Besides this estimation method, we can also use Expectation−Maximization(EM) algorithm to iteratively estimate the hyperparameters, see Xing and Yu (2013).

Chapter 8

An Multiplicative Intensity Model with Smoothly Time-Varying Coefficients

8.1 Model Specification

In Section 2, we focus on the intensity model with unknown structure breaks which assume that the time-varying coefficients are piecewise constant. In other words, the coefficients $\beta(t)$ undergo abrupt change at some time and between the adjacent change-point, $\beta(t)$ is fixed as a constant coefficient. In this section, we will introduce another time-varying setting for the coefficient which assumes that $\beta(t)$ doesn't have abrupt change in the sample period but changes smoothly since we want to compare this two models in the real data analysis. The inference procedures based on smoothing techniques have been developed by Chiang and Wang (2009).

The model is similar with that in last section, we also consider n firms in the study. Let $N_l^\star(t)$ count the total the number of default event that happened on the lth firm ($l = 1, \ldots, n$) over the time window $[0, t]$ and the risk factors $Z_l(\cdot)$ be a p-dimensional risk covariate process of the lth firm. Denoting C_l as censoring time for the lth firm, and we have

$$\mathbb{E}\{\mathrm{d}N_l^\star(t)|Z_l(t), C_l \geq t\} = \mathbb{E}\{\mathrm{d}N_l^\star(t)|Z_l(t)\}, t \geq 0$$

where $\mathrm{d}N_l^\star(t) = \mathrm{d}N_l^\star\{(t+\mathrm{d}t)^-\} - \mathrm{d}N_l^\star(t^-)$, the increment of $\mathrm{d}N_l^\star$ over the small interval $[t, t+\mathrm{d}t)$. For any firm over the period $[0, C]$, the observed data consist of $\{N_l(\cdot), Y_l(\cdot), Z_l(\cdot)\}$ $(l = 1, \ldots, n)$ which is assumed as independent identically distributed. Then we also denote

$$\mathbb{E}\{\mathrm{d}N_l^\star(t)|Z_l(t)\} = \mathrm{d}\mu_l(t) \tag{8.1}$$

and consider the marginal regression model:

$$\mathrm{d}\mu_l(t) = \exp\{\beta(t)^T Z_l(t)\}\mathrm{d}\mu_0(t), \tag{8.2}$$

where $\mu_0(t)$ is still an unknown continuous function. However, different from the setting in section (7.1), the time-varying coefficients $\beta(t) \in \mathcal{R}^d$ are assumed to be a smoothly function which doesn't have abrupt change-point.

8.2 An Estimation Procedure for Smoothly Time-Varying Coefficients $\beta(t)$

The estimation procedure for smoothly time-varying $\beta(t)$ is also based on the estimation equation as we introduced in Section (7.2). As we know that the key idea in smoothing techniques is using the weighted function or kernel $K_d(x_0, x_i)$ to achieve localization. The kernel function $K_d(x_0, x_i)$ assigns a weight to X_i based on its distance from the target x_0. And the kernels K_d are typically indexed by a parameter d that dictates the width of the neighborhood. In our model, we consider one-dimensional Gaussian kernel smoother, that is

$$K_d(t, t_i) = \exp(-\frac{(t_i - t)^2}{2d^2})$$

in which t is the target time spot, and t_i is the neighborhood around the target. Under this setting, the model (7.2) becomes to

$$\mu_l(t) = \int_0^T \exp\{\theta_0^T Z_l(u) K_d(t, u)\} \mathrm{d}\mu_0(u) \tag{8.3}$$

And as the same as we do in Section (7.2), denote

$$S^{(k)}(\theta, t) = n^{-1} \sum_{l=1}^{n} Y_l(t) Z_l(t)^{\otimes k} \exp\{\theta^T Z_l(t)\}, \quad (k = 0, 1, 2)$$

where $t \in (0, T)$. $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$ and $a^{\otimes 2} = aa^T$. And let $\overline{Z}(\theta, t) = S^{(1)}(\theta, t)/S^{(0)}(\theta, t)$, and $\overline{z}(\theta, t)$ be the limit of $\overline{Z}(\theta, t)$. And note that, under the smoothly coefficients assumption, we will obtain the partial likelihood score function for θ is $U^{sm}(\theta, \tau)$, where

$$U_{0,T}^{sm}(\theta, t) = \Sigma^{-1}(\mu - \theta) + \sum_{l=1}^{n} \int_0^t [Z_l(u) - \overline{Z}(\theta, u)] K_d(t, u) \mathrm{d}N_l(u). \tag{8.4}$$

Denote the solution to $U^{sm}(\theta, \tau) = 0$ by $\theta = \widehat{\beta}(\tau)$. To establish the asymptotic distribution of $\widehat{\beta}(\tau)$ under model (8.3), we need to derive the corresponding distribution of $U^{sm}(\theta, \tau)$. Similar with the procedure described in Section (7.2), we note that

$$U_{0,T}^{sm}(\theta_0, t) = \Sigma^{-1}(\mu - \theta_0) + \sum_{l=1}^{n} \int_0^t [Z_l(u) - \overline{Z}(\theta_0, u)] K_d(t, u) \mathrm{d}M_{0,T;l}(u), \tag{8.5}$$

where

$$M_{0,T;l}(t) = \int_0^t \mathrm{d}N_l(u) - \int_0^t Y_l(u) \exp\{\theta_0^T Z_l(t)\} \mathrm{d}\mu_{0,T;0}(u)$$

Similarly, we get derive that the process $n^{-1/2}U^{sm}_{0,T}(\theta, t) \quad (0 \le t \le T)$ converges weakly to a continuous zero-mean Gaussian process with covariance function

$$\Sigma_{0,T}(s,t) = \mathbb{E}[\int_0^s \{Z(u) - \overline{z}(\theta_0, u)\}K_d(s,u)\mathrm{d}M_{0,T}(u) \int_0^t \{Z(v) - \overline{z}(\theta_0, v)\}K_d(t,u)\mathrm{d}M_{0,T}(v)] \tag{8.6}$$

where $0 \le s, t \le T$, and the that $n^{-1/2}(\widehat{\beta}_{(0,T)} - \theta_0)$ is asymptotically zero mean normal vector with covariance matrix

$$\Gamma_{0,T} = A^{-1}_{0,T}\Sigma_{0,T}(0,T)A^{-1}_{0,T} \tag{8.7}$$

in which

$$A_{0,T} = \mathbb{E}[\int_0^T \{Z(t) - \overline{z}(\theta_0, t)\}^{\otimes 2} Y(t) \exp\{\theta_0^T Z(t)\}\mathrm{d}\mu_0(t)]$$

Chapter 9

Real Data Analysis

9.1 Data Description

9.1.1 Firms' Default and Accounting Data

In this study, our data set contains 1818 firms, and the sample period is from *January 1986* to *March 2013*, 327 months, yielding 594486 firms-months of data in total. For each firm, we collect credit rating (default information), accounting information include debt, liability and stock. The detail is presented below:

- **Default** A default is defined as bankruptcy. It's the *event* in our counting process model. The default happens when the firm involves in distressed exchange, dividend omission, grace-period default, indenture modified, missed interest payment, missed principal and interest payments, missed principal payment, payment moratorium or suspension of payments. Our firms' credit rating data is from WRDS. Since the credit rating data is categorized as different rating levels such as "AAA", "BB+" and"CCC". First, we ignore the plus(+) and minus(-) which bring subgroups and combine "CCC", "CC" and "C" as one group since the last two rating categories are relatively few. And then we remove rating records of two invalid ratings "N.M" and "Suspended". After

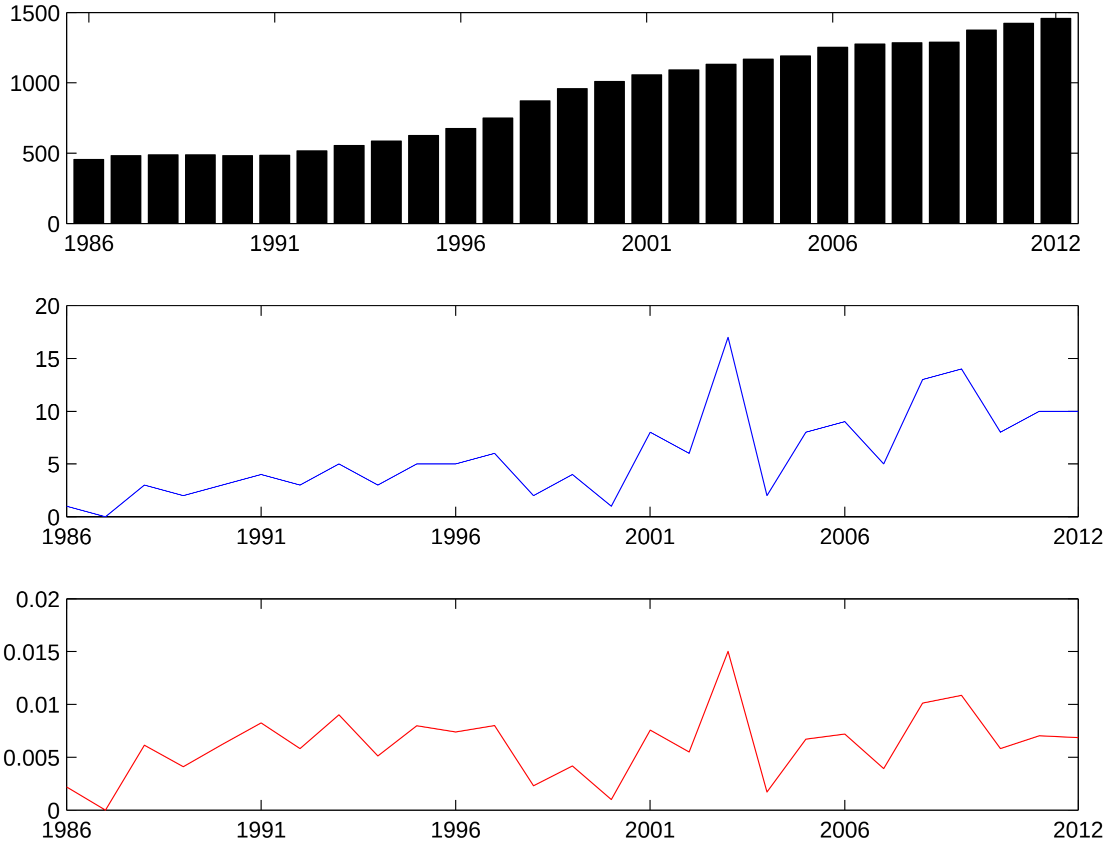

Figure 9.1: Number of survival firms (Upper panel) in each year 1986-2012; Number of default event happened (Middle panel) in each year 1986-2012; The default ratio or default probability during the same period (Bottom panel)

that, we reorganize all ratings into eight groups, that is "AAA", "AA", "A", "BBB", "BB", "B", "C", "D". Secondly, find all the transitions from the first 7 categories to the last "D" category and treat them as default events. When the multiple defaults happens on the same firm, we just write down the last one and ignore all the formers. Then, we truncate default data to fit window 1986-2013. Finally, we get 160 firm default events in this 27-year sample period. As showed in Figure (9.1), the upper bar plot is to display the number of survival firms in each year, while the middle plot is the number of default events happened in each year. In the bottom, we plot the ratio of the two former numbers which represents the default probability during this period. Note that, we also find another 200 companies which exit sometime during this 27-year sample period. The reason for the exit could be acquisitions, mergers, "now a private company" or others. We don't treat these exits event as default.

- **Current Liabilities** A company's debts or obligations that are due within one year. We get *Debt in current Liabilities*(quarterly data) and *Total current Liabilities*(yearly data) from Compustat. Following Moody's KMV (Crosbie and Bohn, 2002), we use build up *Short-term Debt* as the larger one between the two data.

- **Long-term Debt** Long-term debt for a company would include any financing or leasing obligations that are to come due in a greater than 12-month period. Such obligations would include company bond issues or long-term leases that have been capitalized on a firm's balance sheet. This quarterly data is also from Compustat.

- **Equity** The firm's ownership interest which is represented by stock in our study. We collect *Stock price*(monthly, closed price) and *Common Shares Outstanding*(quarterly) from Compustat. Combining them together, we can compute the value of the total equity for the firm. And stock price is also important in the constructing of the risk factor *Firm's trailing one-year stock return.*

- **1-year CMT Year Rate** The Constant Maturity Treasury rate is interpolated one-year yield of the most recently auctioned 4-, 13- and 26-week U.S. Treasury bills, plus the most recently auctioned 2-, 3-, 5- and 10-year U.S. Treasury notes as well as the most recently auctioned U.S. Treasury 30-year bond, plus the off-the-runs in the 20-year maturity range. This monthly data is obtained from Federal Reserve Bank of St. Louis.

- **CRSP NYSE/AMEX index return** It's CRSP Value-weighted index returns which combine NYSE and AMEX. The monthly data is from WRDS. It's important and like a benchmark in constructing the *Firm's trailing one-year stock return*.

9.1.2 Covariates

The default risk factors or covariates used in our study are the same as that used to estimates the models of Duffie et al.(2007), Das et al.(2007) and Duffie et al.(2009). Before building up the covariates, we first transfer all the quarterly and yearly data to monthly data by using linear interpolation, which is the preparation for the monthly covariates construction.

- **Distance to Default** The firm's *distance to default* is the number of standard deviations of asset growth by which assets exceed a standardized measure of liabilities. It's a volatility-adjusted measure of leverage. The construction method we adopted here is that used by Vassalou and Xing (2004), Crosbie and Bohn (2002), and Hillegeist et al. (2004). Although the conventional approach to measure distance to default involves some rough approximations, Bharath and Shumway (2008) provide evidence that default prediction is relatively robust to varying the proposed measure with some relatively simple alternatives. Let talk about the construction method in detail.

The *distance to default* D_t can be expressed and calculated by

$$D_t = \frac{\ln(V_t/L_t) + (\mu_A - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}}, \tag{9.1}$$

in which

V_t: Market value of the firm's assets at time t.

L_t: Liability.

μ_A: Firm's mean rate of asset growth.

σ_A: Firm's asset volatility.

Computation Steps:

L_t: By Moodys KMV (Crosbie & Bohn, 2002), L_t, the Liability, equals to "short term debt + 0.5 × long term debt" where *short term debt* is max of *Debt in current Liability* and *Total current Liability.*

σ_A and V_t: By Merton(1974) and Black and Scholes(1973), we can calculate V_t and σ_A by iteratively applying the equations:

$$\begin{cases} W_t &= V_t\Phi(d_1) - L_t e^{-r_t T}\Phi(d_2) \\ \sigma_A &= std(\ln(V_t) - \ln(V_t - 1)) \end{cases}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function and

$$d_1 = \frac{\ln(V_t/L_t) + (r - \frac{1}{2}\sigma_A^2)T}{\sigma_A\sqrt{T}}$$

$d_2 = d_1 - \sigma_A\sqrt{T}$

W_t = stock price $\times$ number of shares outstanding + total debt (short term debt + long term debt). Taking the initial asset value V_t to be the sum of W_t and risk-free return r_t to be the 1-year CMT Year Rate. The iteration stops when the σ_A converges.

μ_A: Following Vasslou & Xing (2004), we calculate μ_t vector by

$$\mu_t = \max(\frac{V_t - V_{t-1}}{V_{t-1}}, r_t)$$

in which r_t is the 1-year CMT Year Rate. And then we can obtain μ_A as $mean(\max(\log \frac{V_t}{V_{t-1}}, r_t))$ when V_t is known.

Plug all the results into (9.1), then we build up the *distance to default*.

- **Trailing 1-year stock return** The firm's *trailing 1-year stock return* or firm's *past excess stock return* is an important default risk factor suggested by Shumway (2001). We do not have a particular structural interpretation for this covariate, however, Shumway (2001) and Duffie et al. (2009) have found that this covariate offers significant incremental explanatory power, perhaps as a proxy for some unobserved factor that has an influence on default risk beyond that of the firm's measured distance of default. We follow Shumway (2001) to calculate it by the difference of firm's 1-year stock return and CRSP NYSE/AMEX index return.

9.2 Model Estimation under Structure Breaks Assumption

9.2.1 Time-varying $\beta(t)$

In this dissertation research, our interest is to analyze and understand the relationship of risk factors and default events under market structural break. As we introduce in last section, to link the intensity of the counting process $N^\star(t)$ with firm's risk factors, we consider two covariates in the study: firm's *distance to default* and firm's *trailing 1-year stock return.*

In the 1818 firms, there are 160 defaults happened in the sample period from January 1986 to March 2013. Applying the time-varying $\beta(t)$ estimation procedure we present in Section (7.4)-(7.6) with the two covariates which are constructed from firm's specific accounting information and macroeconomic data, we can calculate calculate the time-varying $\beta_1(t)$ and $\beta_2(t)$(see Figure 9.2 and Figure 9.3) which are the coefficients for *distance to default* and *trailing stock return*, respectively.

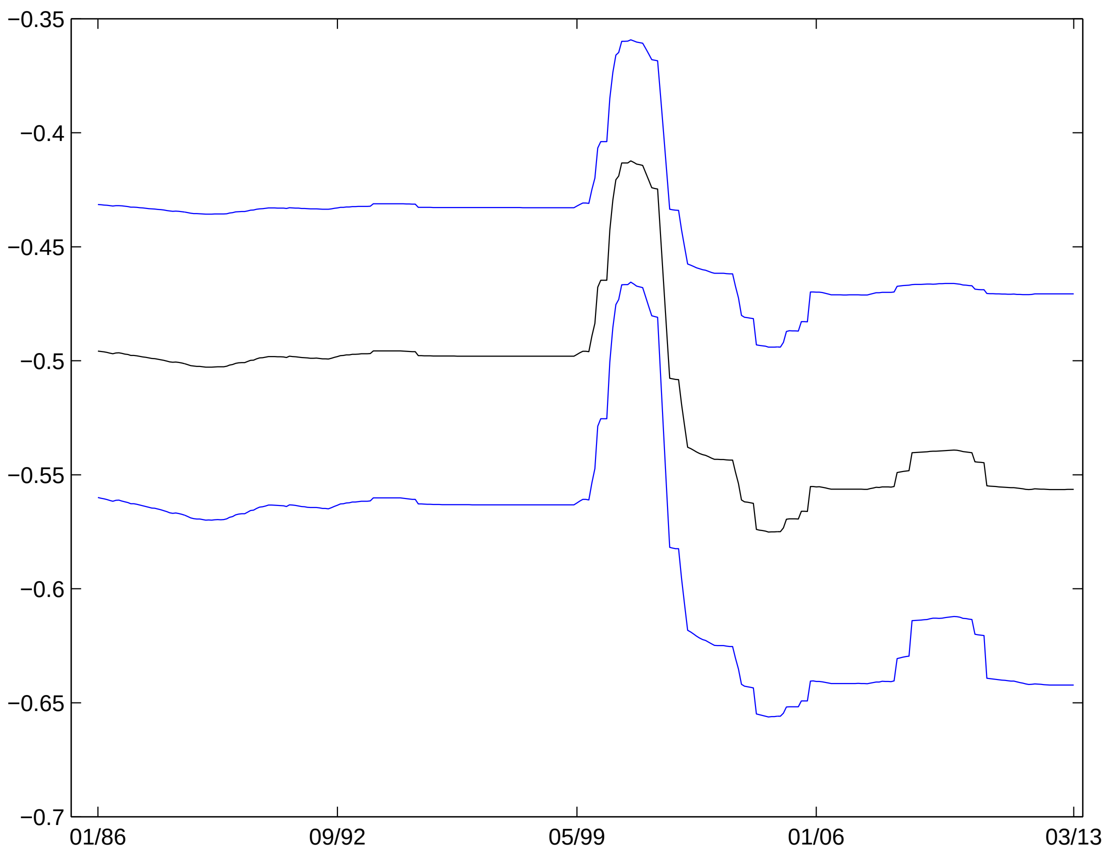

Figure 9.2: Time-varying $\beta_1(t)$. The middle black line is the $\widehat{\beta}_1(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation.

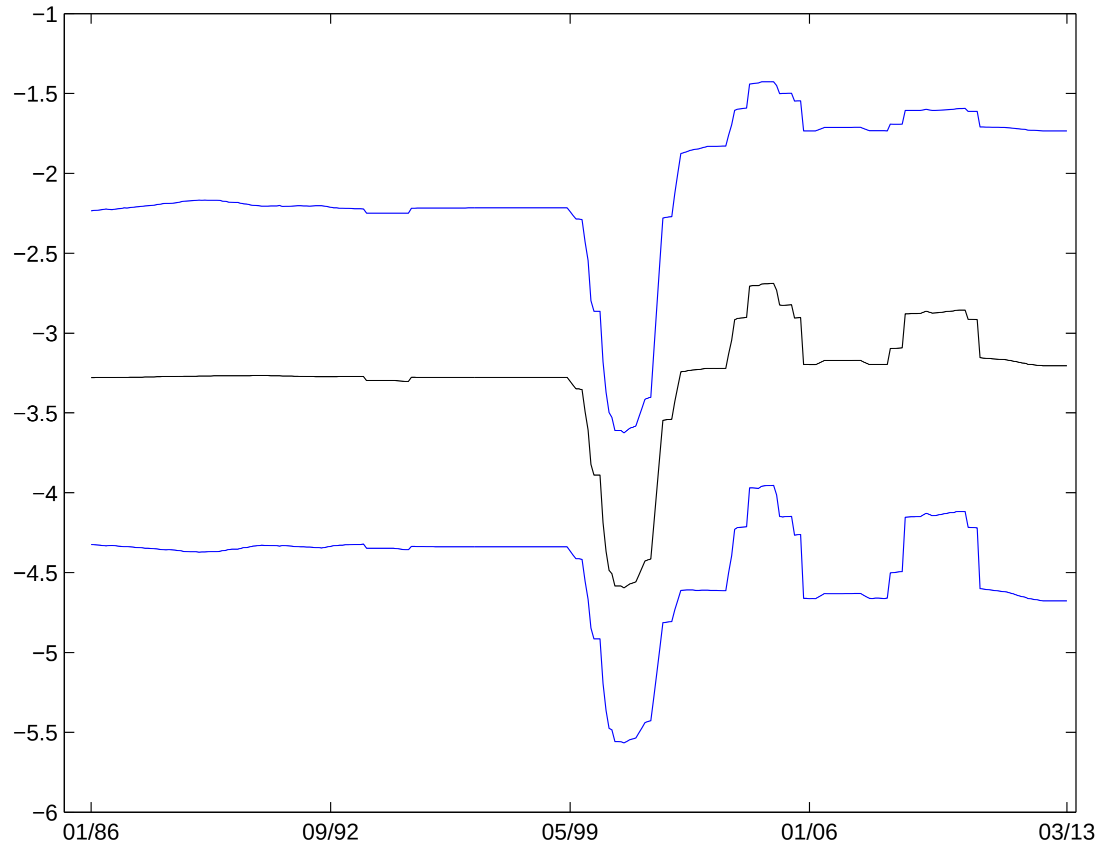

Figure 9.3: Time-varying $\beta_2(t)$. The middle black line is the $\widehat{\beta}_2(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation.

As shown in Figure (9.2) and (9.3), the distance to default and trailing stock return have negative coefficients for the default intensity which matches the the information represented by the two risk factors. The company which has longer or larger distance to default and higher trailing stock return will be unlike to default or have a lower default intensity.

From the plot (9.2) and (9.3), we can find the coefficients went through a stable or constant period from 1986 to 1999. However, we know the fact that in October 1987, US stock market economy crashed, shedding a huge value in a very short time. And from 1990-1991, a recession hit the economy but after March 1991, the economy began to recovery (Announced by NBER). We think the coefficients kept stable in the changing economy environment because the stock market crisis and short term recession didn't have large impact on the firms' default. There's less than 5 default events happened averagely in the period which can also support our judgment.

Between May 1999 and October 2002, the coefficients have an abrupt change. The coefficient of distance to default went up first then went down, in contrast, the coefficient of trailing stock return went down first then went up. The same abrupt change in time but different in style is very interesting. First, we know that from the second half of 1998, a series of devastating events happened to give large impact to the economy not just to stock market, including Russia's default, Brazil's currency crisis and sever disruption of LTCM(Long-Term Capital Management L.P.). On March 10, 2000, NASDAQ composite index peaked at 5,048.62, more than double its value just a year before, however, the Dot-com bubble began to burst. On March 20, 2000, NASDAQ has lost more than 10% from its peak. By 2001 the bubble was deflating at full speed. The series of events bring a storm to U.S. company which make a lot of firm default and the risk factors' coefficients change a lot in short time. The different movement direction of β_1 and β_2 may have such a reason: when the bad economy time come, the performance in stock market becomes more important (β_2

increases in absolute value)as a default risk factor, while the distance to default has weaker (β_1 decreases in absolute value) influence. And vice versa when economy began to recover.

The second abrupt change in coefficients happened between November 2003 and October 2005. To stimulate economy, Fed lowered the federal funds rate after the Dot-com bubble burst, till May, 2003, the federal funds rate was down to the 40-year lowest point, turning the US economy from recovery to healthy expansion.

The last notable unstable period for the coefficients is from June 2008 to September 2010. It's not surprise to find that the financial crisis happened from 2008 makes the default risk factors' coefficient break constant situation and go through a turmoil. Both β_1 and β_2 become larger in this period, or decrease in their absolute value. That mean, the risk factors has weaker impact to the default event. We guess the reason for the phenomenon is the firms' or investors' tolerance to bad financial position becomes stronger when facing the worst financial crisis since the Great Depression of the 1930s .

9.2.2 Baseline Intensity

Recall our model 7.3

$$\mu_l(t) = \int_0^t \exp\{\beta(u)^T Z_l(u)\} \mathrm{d}\mu_0(u)$$

where $\mu_0(t)$ is the baseline hazard function or baseline intensity which is assumed as an unknown and unspecified continuous function. This baseline function represents the common macroeconomic risk factor. After we obtain the time-varying coefficients $\beta(t)$, we can calculate $\widehat{\mu}_0(t)$the baseline function by the Aalen-Breslow type estimator which is showed in formula (7.10). To explore it, we plot the baseline $\mu_0(t)$ in Figure (9.4). From the plot, we can find that the baseline function stay almost constant from 1986 to 1999. There are only

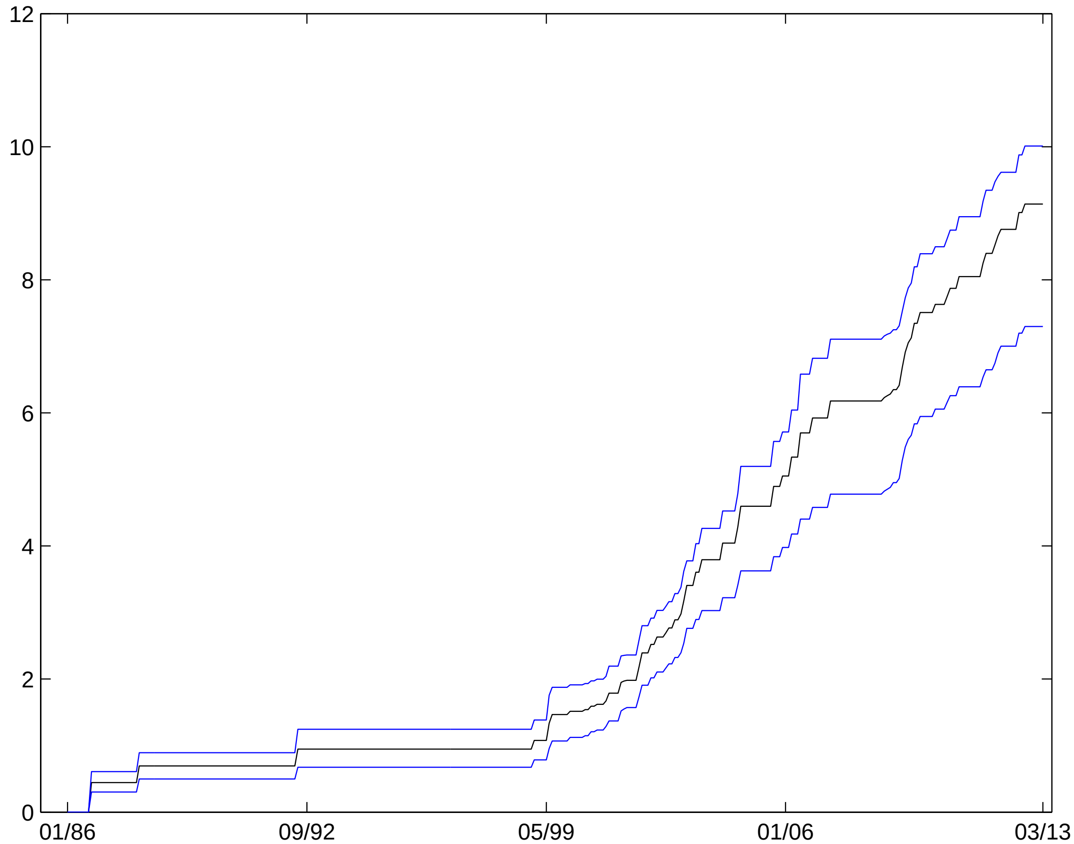

Figure 9.4: The baseline function $\mu_0(t)$. The middle black line is the $\widehat{\mu}_0(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation.

two months which manifest fast increase: January 1988 and June 1992. However, after May 1999, the baseline function enters into a "non-stable" track. Between these abrupt increases, the periods from December 2001 to August 2003 and from September 2008 to October 2009 is the fastest. That means the macroeconomic risk factors become more hazardous to the companies in our sample. Look at these two periods, one is just in the full speed deflation of Dot-com bubble and the other is in the worst time of the financial crisis. A more interesting finding is that the fastest increase in baseline function in a single month happens in May 2006 and this month is considered as the start point of US housing prices rise before financial crisis and finally leads to it.

9.2.3 Firm's Default Intensity

Now, let's take a look at the company's intensity function $\mu_l(t)$. In this analysis, we pick AMR Corpoaration as our individual firm sample. AMR Corporation is a commercial aviation business and airline holding company based in Fort Worth, Texas. And it's best known for being the parent company of American Airlines. The company also owns AMR Eagle Holdings Corporation, which operates the regional airlines American Eagle Airlines and Executive Airlines. We can easily calculate AMR's intensity function when we have the time-varying coefficients and baseline function as showed as formula (7.3) and plot it in Figure (9.5).

$$\mu_{AMR}(t) = \int_0^t \exp\{\beta(u)^T Z_{AMR}(u)\} \mathrm{d}\mu_0(u)$$

From the plot, we can notice that the first large increase in the intensity function happens in the period from 2002 to 2003. Check the company's business record, we know that in response to decreased demand following the events of September 11, 2001, the company reduced its operating schedule by approximately 20 percent and reduced its workforce by approximately 20,000 jobs. And in August 2002, the company was cutting another 7,000 jobs. First-class

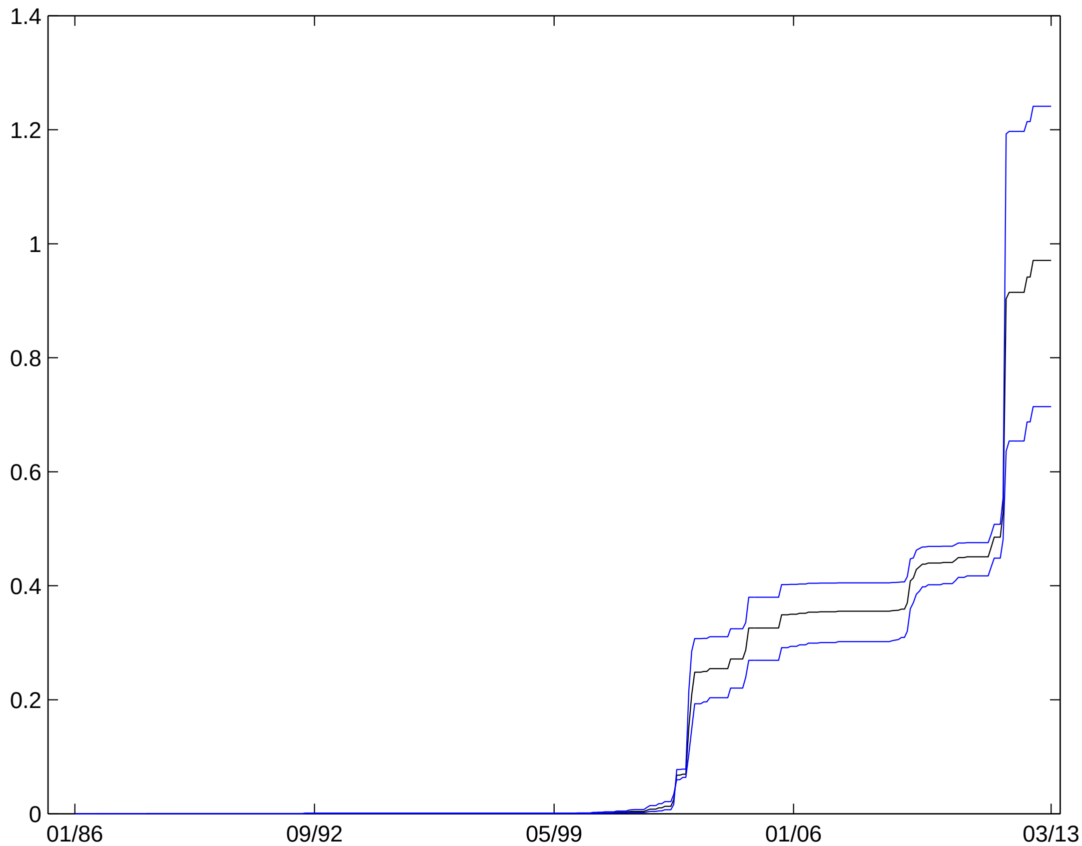

Figure 9.5: The AMR Corp. Intensity function $\mu_{AMR}(t)$. The middle black line is the $\widehat{\mu}_{AMR}(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation.

service was removed on most flights, with the exception of major international routes. Even the company reduced its annual costs by \$2 billion in 2002, it still incurred net losses of \$3.5 billion, the worst year in the company's history. However, after second quarter of 2003, in conjunction with the improvement of US economy, the company's revenue environment began to improve. In the plot, we can see the intensity enters into a relatively stable period after 2003.

In financial crisis period, the company also encountered the operating difficulties. From the Figure (9.5), it's obvious that find the fast increase in intensity function $\mu_{AMR}(t)$ in the period. The most recent event happens on AMR is in November 2011. The company filed for Chapter 11 reorganization bankruptcy with \$4 billion of cash. The worse is in February 2012, the company announced that in order to cut operating costs and boost revenue, it would eliminate 13,000 jobs, which amounted to 18 percent (including 15 percent management positions) of American Airline's 73,800 employees. And in 2012, AMR began to consider merge with US Airways. In the plot, we can also find this large intensity function increase period starting from Nov, 2011. At the same time, the estimates variance for the intensity function also becomes much larger than before which may reveal the uncertainty of business situation of the company.

9.3 Model Estimation under Smoothly Time-Varying Coefficients Assumption

As a comparison with the model which is under the smoothly time-varying coefficients assumption, we plug the real data into the model described in Chapter (8). We calculate the smoothly time-varying $\beta_1(t)$ and $\beta_2(t)$, and with estimated coefficients, we then to compute the baseline function $\mu_0(t)$. They're showed in Figure (9.6), Figure (9.7) and Figure (9.8), respectively. Meantime all the 95% confidence intervals will also be provided.

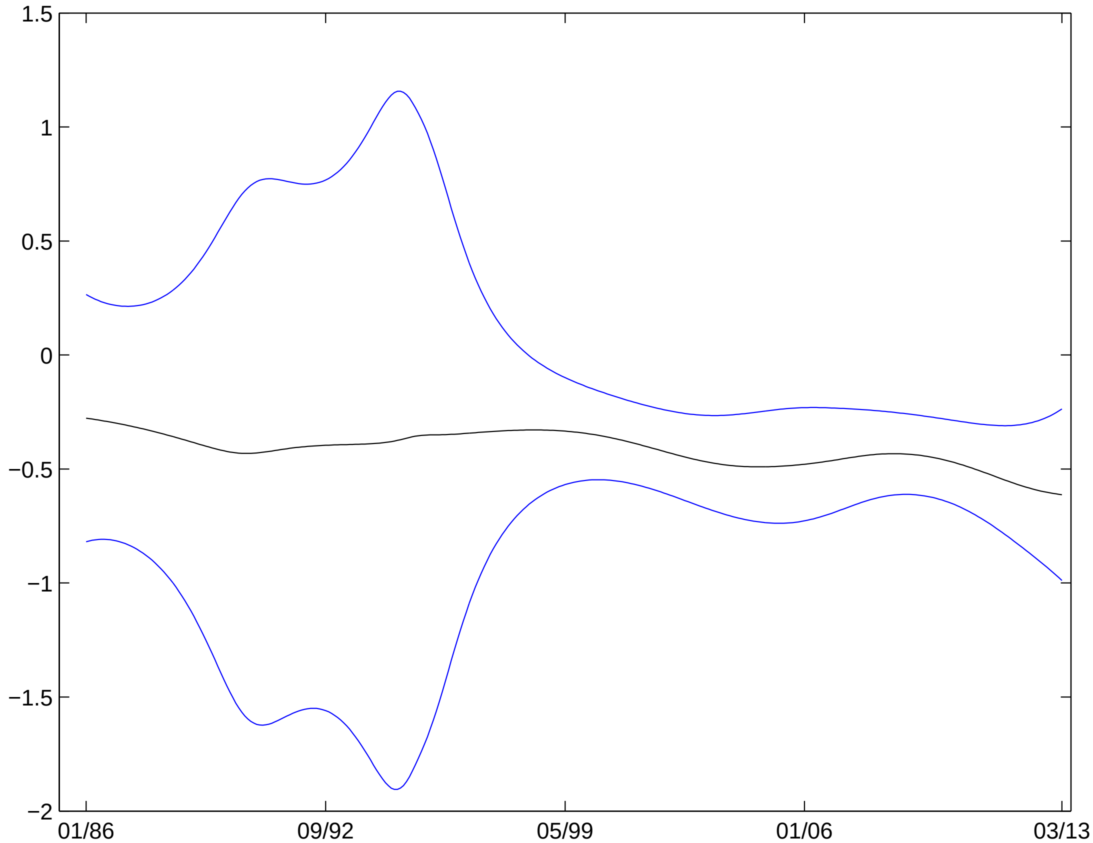

Figure 9.6: Smoothly time-varying $\beta_1(t)$. The middle black line is the $\widehat{\beta}_1(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation.

From the Figure (9.6) and Figure (9.7), we can clearly find two common features which are very different from the results based on our model. First, the estimation of the two risk factors' coefficients is relatively stable and rarely has abrupt up or downs. Even this result fits the assumption of smoothing function setting, it fails to capture any sudden changes in the coefficients. Second, for the early period, the estimation provided by smoothing model incurs large variance, leading the 95% confidence interval to contain 0 in it. From the plots, we can tell that such large confidence interval actually make the estimation very weak in explanatory power. And the same feature can also be found from the Figure (9.8):

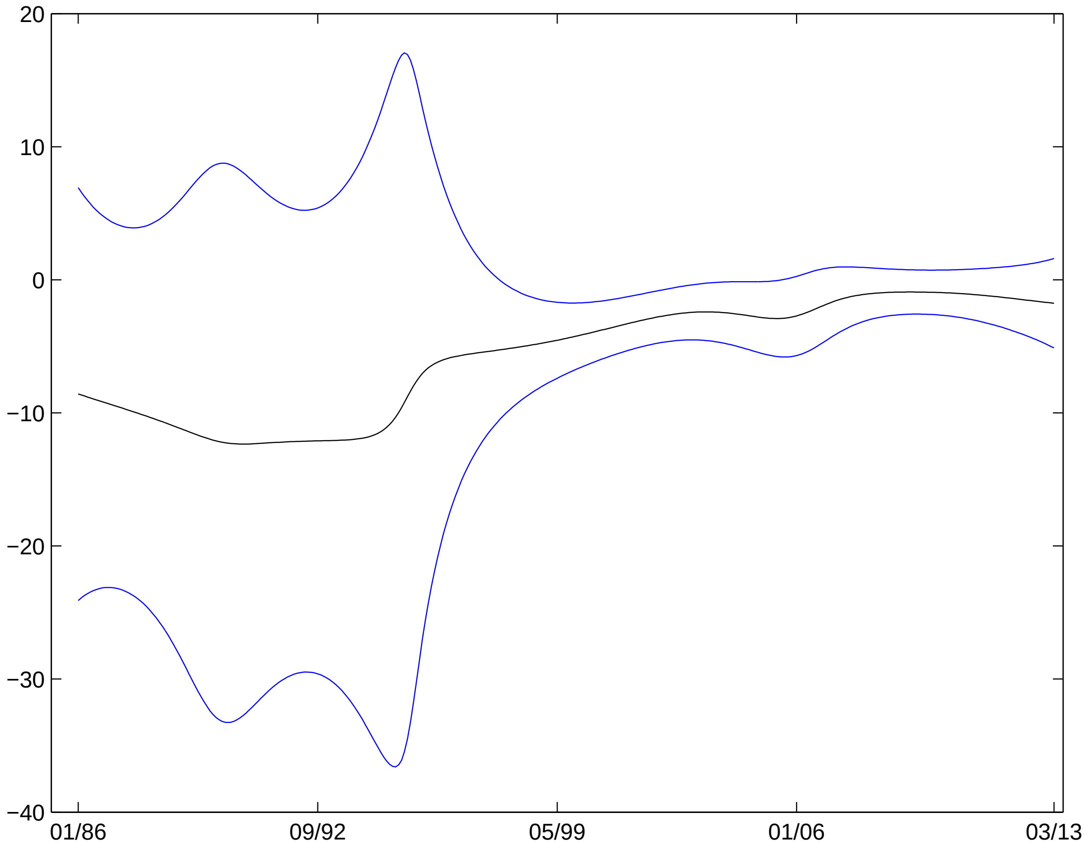

Figure 9.7: Smoothly time-varying $\beta_2(t)$. The middle black line is the $\widehat{\beta}_2(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation.

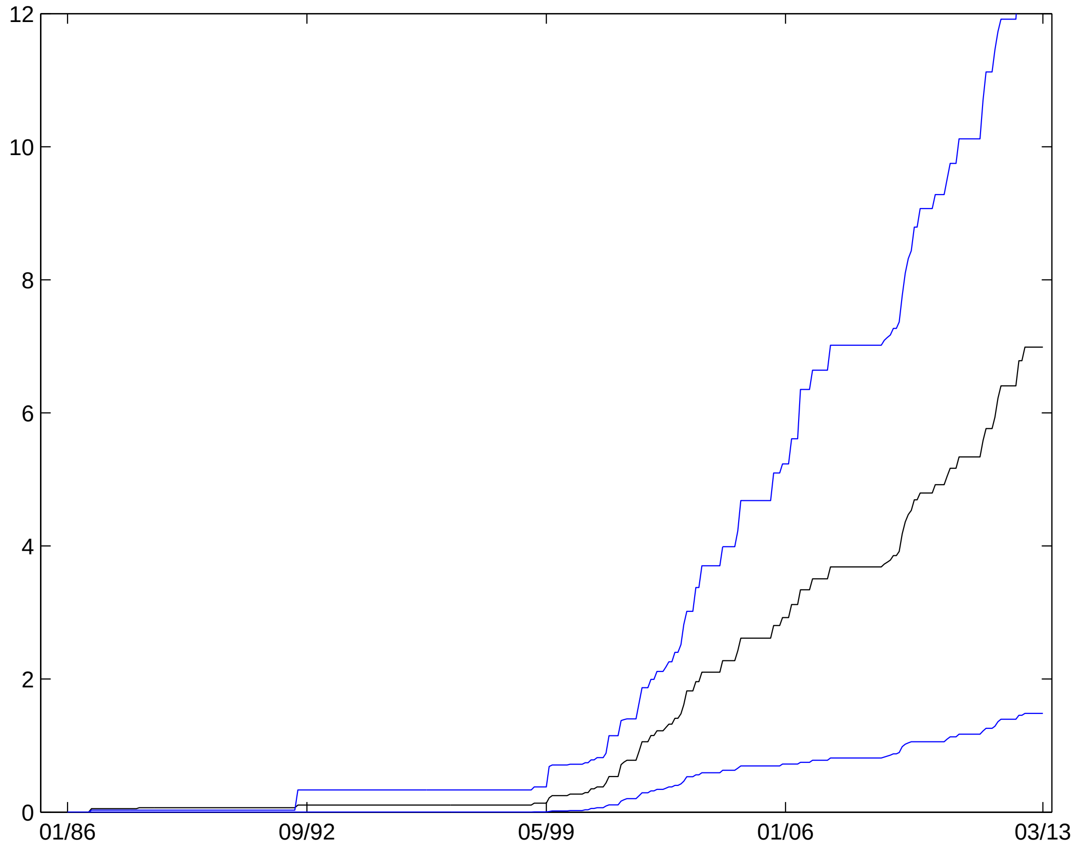

Figure 9.8: The baseline function $\mu_0(t)$. The middle black line is the $\widehat{\mu}_0(t)$ and two blue line confine the pointwise 95% confidence interval for the estimation

this estimation fails to show the fast increase of baseline function in short time and the much worse thing is that the variance of the estimation coming from smoothing model is too large compared to our model's. From this comparison, we have seen the incompetent of the smoothing time-varying coefficients model on dealing with the abrupt change-points environment.

Chapter 10

Concluding Remarks

This dissertation research is focused on the dynamics of firms' default risk when the risk factors' coefficients have structure changes. We consider an advanced Cox type semiparametric model in survival analysis to modulate firm's default intensity with two default risk factors or covariates *distance to default* and *firm's trailing stock return*. For firm l and at time t, the model is:

$$\mathrm{d}\mu_l(t) = \exp\{\beta(t)^T Z_l(t)\}\mathrm{d}\mu_0(t),$$

We assume the $\beta(t)$ is piecewise constant and have unknown number and unknown time change-point. And we also construct an estimation procedure for the $\beta(t)$ and its asymptotic covariance based on the the estimating equation for the constant β in previous model. This estimation procedure outputs the posterior distribution of the coefficients as a mixture distribution of the segmented constant coefficients with explicit weights which can be calculated recursively using a Bayesian method. In addition, to lower the computational complexity, a Bounded Complexity Mixture Approximation (BCMIX) is presented. Compared to the Bayes estimates, the BCMIX has higher efficiency. To prepare the real data analysis, we also introduced a previous model which assumes the smoothly time-varying coefficients, including this model's estimation procedure.

In the real data analysis, we collect the default and accounting data of 1818 firms from *January 1986* to *March 2013*, 327 months, yielding 594486 firms-months of data in total. Then we calculate the time-varying coefficients with structural breaks and baseline intensity function and we discuss how the results capture the real events happened in history. We pick AMR Corporate as the individual firm in the analysis and compute its default intensity as an example. We also plot and talk about the results from the model which has smoothly time-varying coefficients.

With our model, we successfully solve the problem of estimating the time-varying default risk factors' coefficients in the presence of structural breaks. We can derive analytical filtering formulas for the posterior distributions of the parameters and the approximation algorithm BCMIX lower the computational complexity and improve the statistical efficiency.

Reference

Ait-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. *Journal of Finance*, 57, 2075-2112.

AitSahlia, F. and Lai, T.L. (1999). A canonical optimal stopping problem for American options and its numerical solution. *Journal of Computational Finance*, 3(2), 33-52.

AitSahlia, F. and Runnemo, A. (2007). A canonical optimal stopping problem for American options under a double exponential jump-diffusion model. *Journal of Risk*, 10, 85-100.

Altman, E.I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4), 589-609.

Andersen, E.D. and Damgaard, A. (1999). Utility based option pricing with proportional transaction costs and diversification problems: an interior-point optimization approach. *Applied Numerical Mathematics*, 29, 395-422.

Andersen, P.K., and Gill, R.D. (1982). Cox's regression model for counting processes: a large sample study. *The annals of statistics*, 1100-1120.

Andersen, T. and Benzoni, L. and Lund, J. (2002). An empirical investigation of continuous-time equity return models. *Journal of Finance*, 57(3), 1239-1284.

Anderson, R.W., and Sundaresan, S. (1996). Design and valuation of debt contracts. *Review of financial studies*, 9(1), 37-68.

Bachelier, L. (1900). Théorie de la spéculation. *Annales de l'Ecole Normale Supérieure*, 17, 21-86.

Banachewicz, K., and Lucas, A. (2008). Quantile forecasting for credit risk management using possibly misspecified hidden markov models. *Journal of Forecasting*, 27(7), 566-586.

Bather, J.A. and Chernoff, H. (1966). Sequential decisions in the control of a spaceship. *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 3, 181-207.

Beaver, W.H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, 71-111.

Beaver, W.H. (1968). The information content of annual earnings announcements. *Journal of accounting research*, 67-92.

Bensaid, B., Lesne, J.P., Pages, H., and Scheinkman, J. (1992). Derivative asset pricing with transaction costs. *Mathematical Finance*, 2, 63-86.

Bharath, S.T., and Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies*, 21(3), 1339-1369.

Bielecki, T.R., and Rutkowski, M. (2002). *Credit risk: modeling, valuation and hedging*, Springer.

Black, F. and Cox, J.C. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *The Journal of Finance*, 31(2), 351-367.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637-659.

Boetius, F. (2005). Bounded variation singular stochastic control and dynkin game. *SIAM J. Control Optim.*, 44(4), 1289-1321.

Bole, P.P. and Vorst, T. (1992). Option replication in discrete time with transaction costs. *Journal of Finance*, 47, 271-293.

Bouchard, B., Kabanov, Y., and Touzi, N. (2001). Option pricing by large risk aversion utility under transaction costs. *Decisions in Economics and Finance*, 24, 127-136.

Broadie, M., Chernov, M., and Johannes, M. (2007). Model specification and risk premia:Evidence from future options. *Journal of Finance*, 62, 1453-1490.

Campbell, J.Y., Hilscher, J., and Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, 63(6), 2899-2939.

Carling, K., Jacobson, T., Lind, J., and Roszbach, K. (2007). Corporate credit risk modeling and the macroeconomy. *Journal of Banking & Finance*, 31(3), 845-868.

Carr, P. and Wu, L. (2003). What type of process underlies options? a simple robust test. *Journal of Finance*, 58, 2581-2610.

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

Chernoff, H. (1961). Sequential test for the mean of a normal distribution. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 79-91.

Chernoff, H. and Petkau, A.J. (1986). Numerical solutions for bayes sequential decision problems. *SIAM Journal on Scientific and Statistical Computing*, 7, 46-59.

Chiang, C.T. and Wang, M.C. (2009). Varying-coefficient model for the occurrence rate function of recurrent events. *Ann.Inst. Stat. Math.*, 61, 197-213

Clewlow, L. and Hodges, S. (1997). Optimal delta-hedging under transaction costs. *Journal of Economic Dynamics and Control*, 21, 1353-1376.

Collin-Dufresne, P., and Goldstein, R.S. (2001). Do credit spreads reflect stationary leverage ratios?. *The Journal of Finance*, 56(5), 1929-1957.

CollinDufresne, P., Goldstein, R., and Hugonnier, J. (2004). A general formula for valuing defaultable securities. *Econometrica*, 72(5), 1377-1407.

Collin-Dufresne, P., Goldstein, R.S., and Helwege, J. (2010). Is Credit Event Risk Priced? Modeling Contagion via the Updating of Beliefs (No. w15733). *National Bureau of Economic Research*.

Constantinides, G.M. and Zariphopoulou, T. (1999). Bounds on prices contingent claims in an intertemporal economy with proportional transaction costs and general preferences. *Finance and Stochastic*, 3, 345-369.

Constantinides, G.M. and Zariphopoulou, T. (2001). Bounds on prices of contingent claims in an intertemporal setting with proportional transaction costs and multiple securities. *Mathematical Finance*, 11, 331-346.

Cont, R.and Tankov, P.(2004). *Financial modelling with jump processes*, Chapman & Hall/CRC.

Crosbie, P.J., and Bohn, J.R. (2002). *Modeling Default Risk*, KMV, San Francisco.

Dai, Q., and Singleton, K.J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, 55(5), 1943-1978.

Das, S. R., Duffie, D., Kapadia, N., and Saita, L. (2007). Common failings: How corporate defaults are correlated. *The Journal of Finance*, 62(1), 93-117.

Davis, M. and Lo, V. (2001). Infectious defaults. *Quantitative Finance*, 1, 382-387.

Davis, M.H.A. and Norman, A.R. (1990). Portfolio selection with transaction costs. *Mathematics of Operation Research*, 15, 676-713.

Davis, M.H.A., Panas, V.G., and Zariphopoulou, T. (1993). European option pricing with transaction costs. *SIAM J. Control Optim.*, 31(2), 470-493.

Derman, E. and Kani, I. (1994). *Riding on a smile*, RISK, 7, 32-39.

Duffie, D., Eckner, A., Horel, G., and Saita, L. (2009). Frailty correlated default. *The Journal of Finance*, 64(5), 2089-2123.

Duffie, D., and Gârleanu, N. (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 41-59.

Duffie, D., and Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica*, 69(3), 633-664.

Duffie, D., and Liu, J. (2001). Floating-fixed credit spreads. *Financial Analysts Journal*, 76-87.

Duffie, D., and Singleton, K.J. (1999). Modeling term structures of defaultable bonds. *Review of Financial studies*, 12(4), 687-720.

Duffie, D., Saita, L., and Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. *Journal of Financial Economics*, 83(3), 635-665.

Dupire, B. (1994). *Pricing with a smile*, RISK, 7, 18-20.

Edirisinghe, C., Naik, V., and Uppal, R. (1993). Optimal replication of options with transaction costs and trading restrictions. *Journal of Finance and Quantitative Analysis*, 28, 117-138.

Eraker, B., Johannes, M.S., and Poison, N. (2003). The impact of jumps in equity index volatility and returns. *Journal of Finance*, 58, 1269-1300.

Feng, D., Gourieroux, C., and Jasiak, J. (2008). The ordered qualitative model for credit rating transitions. *Journal of Empirical Finance*, 15(1), 111-130.

Fisher, E. Heinkel R. and Zechner J.(1989). *Dynamic Capital Structure Choice: Theory and Tests*, 19-40.

Frey, R., and McNeil, A.J. (2003). Dependent defaults in models of portfolio credit risk. *Journal of Risk*, 6, 59-92.

Frydman, H., and Schuermann, T. (2008). Credit rating dynamics and Markov mixture models. *Journal of Banking & Finance*, 32(6), 1062-1075.

Fuertes, A.M., and Kalotychou, E. (2007). On sovereign credit migration: A study of alternative estimators and rating dynamics. *Computational Statistics & Data Analysis*, 51(7), 3448-3469.

Geske, R. (1977). The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis*, 12(04), 541-552.

Giampieri, G., Davis, M., and Crowder, M. (2005). Analysis of default data using hidden Markov models. *Quantitative Finance*, 5(1), 27-34.

Giesecke, K. (2004). Correlated default with incomplete information. *Journal of Banking & Finance*, 28(7), 1521-1545.

Gordy, M.B. (2000). A comparative anatomy of credit risk models. *Journal of Banking & Finance*, 24(1), 119-149.

Grinold, R.C. and Kahn, R.N. (2000). *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*, Second ed., McGraw-Hill, New York.

Heston, S. (1993). A closed-form solution of options with stochastic volatility with applications to bond and currency options. *Review of Financial Study*, 6, 327-343.

Hillegeist, S.A., Keating, E.K., Cram, D.P., and Lundstedt, K.G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies*, 9(1), 5-34.

Hodges, S.D. and Neubcrger, A. (1989). Optimal replication of contingent claims under transactions costs. *Review of Futures Markets*, 8, 222-239.

Hui, C.H., Lo, C.F., and Huang, M.X. (2006). Are corporates' target leverage ratios time-dependent?. *International Review of Financial Analysis*, 15(3), 220-236.

Hui, C.H., Lo, C.F., and Tsang, S.W. (2003). Pricing corporate bonds with dynamic default barriers. *Journal of Risk*, 5(3), 17-37.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, pages 281-300.

Jarrow, R.A., and Turnbull, S.M. (1995). Pricing derivatives on financial securities subject to credit risk. *The journal of finance*, 50(1), 53-85.

Jarrow, R.A., Lando, D., and Turnbull, S.M. (1997). A Markov model for the term structure of credit risk spreads. *Review of Financial studies*, 10(2), 481-523.

Jorion, P., and Zhang, G. (2007). Good and bad credit contagion: Evidence from credit default swaps. *Journal of Financial Economics*, 84(3), 860-883.

Kadam, A., and Lenk, P. (2008). Bayesian inference for issuer heterogeneity in credit ratings migration. *Journal of Banking & Finance*, 32(10), 2267-2274.

Kalbfleisch, J.D., and Prentice, R.L. (2002). *The statistical analysis of failure time data. Second Ed.*, Wiley, Hoboken, NJ

Karatzas, I. (1983). A class of singular stochastic control problems. *Applied Probability Trust*, 15(2), 225-254.

Karatzas, I. and Shreve, S.E. (1985). Connection between optimal stopping and singular stochastic control ii. reflected follower problems. *SIAM J. Control Optim.*, 23(3), 433-451.

Kealhofer, S. (2003). Quantifying credit risk I: default prediction. *Financial Analysts Journal*, 30-44.

Kou, S.G. (2002). A jump diffusion model for option pricing. *Management Science*, 48(8), 1086-1101.

Kushner, H.J. and Dupuis, P.G. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York.

Kusuoka, S. (1999). A remark on default risk models. *In Advances in mathematical economics*, (pp. 69-82), Springer Japan.

Lai, T.L. and Lim, T.W. (2002). A new approach to singular stochastic control in optimal hedging and investment-consumption under transaction costs. *Working paper*

Lai, T.L. and Lim, T.W. (2009). Option hedging theory under transaction costs. *Journal of Economic Dynamics and Control*, 33(12), 1945-1961.

Lai, T.L. and Wong, S.P.S. (2004). Valuation of American options via basis functions. *Automatic Control, IEEE Transactions on*, 49(3), 374-385.

Lai, T.L. and Xing, H. (2011). A simple Bayesian approach to multiple change-points. *Statistica Sinica*, 21(2), 539.

Lai, T.L., Yao, Y.C., and AitSahlia, F. (2007). Corrected random walk approximations to free boundary problems in optimal stopping. *Advances in Applied Probability*, 39, 753-775.

Lancaster, T. (1992). The econometric analysis of transition data (No. 17). *Cambridge university press.*

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives research*, 2(2-3), 99-120.

Lando, D. (2009). *Credit risk modeling: theory and applications*, Princeton University Press.

Lane, W.R., Looney, S.W., and Wansley, J.W. (1986). An application of the Cox proportional hazards model to bank failure. *Journal of Banking and Finance*, 10(4), 511-531.

Lawless, J.F., and Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics*, 37(2), 158-168.

Lawless, J.F., Nadeau, C., and Cook, R.J. (1997). Analysis of mean and rate functions for recurrent events. *In Proceedings of the First Seattle Symposium in Biostatistics*, (37-49), Springer US.

Lee, S.H., and Urrutia, J.L. (1996). Analysis and prediction of insolvency in the property-liability insurance industry: a comparison of logit and hazard models. *Journal of Risk and Insurance*, 121-130.

Leland, H.E. (1985). Option pricing and replication with transaction costs. *Journal of Finance*, 40, 1283-1301.

Leland, H.E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *The journal of finance*, 49(4), 1213-1252.

Leland, H.E., and Toft, K.B. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 51(3), 987-1019.

Li, D.X. (2000). On default correlation: a copula function approach. *The Journal of Fixed Income*, 9(4), 43-54.

Lin, D.Y., Wei, L.J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4), 711-730.

Longstaff, F.A., and Schwartz, E.S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3), 789-819.

Lucas, A., and Klaassen, P. (2006). Discrete versus continuous state switching models for portfolio credit risk. *Journal of Banking & Finance*, 30(1), 23-35.

Madan, D. and Seneta, E. (1990). The variance-gamma (v.g.) model for share market returns. *Journal of Business*, pages 511-524.

Madan, D.B., and Unal, H. (1998). Pricing the risks of default. *Review of Derivatives Research*, 2(2-3), 121-160.

McDonald, C.G., and Van de Gucht, L.M. (1999). High-yield bond default and call risks. *Review of Economics and Statistics*, 81(3), 409-419.

Merton, R.C. (1973). Theory of rational option pricing. *BELL Journal of Economics*, 4, 141-183.

Merton, R.C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2), 449-470.

Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economic*, 3, 125-144.

Monteiro, A., Lucas, A., and Smirnov, G. (2006). Nonparametric estimation for non-homogeneous semi-Markov processes: an application to credit risk. *Tinbergen Institute Discussion Paper*.

Nickell, P., Perraudin, W., and Varotto, S. (2000). Stability of rating transitions. *Journal of Banking & Finance*, 24(1), 203-227.

Ohlson, J.A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 18(1), 109-131.

Pepe, M.S., and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association*, 88(423), 811-820.

Peskir, G. and Shiryaev, A. (2006). *Optimal Stopping and Free-Boundary problems*, Birkhauser Verlag, Basel-Boston-Berlin.

Pollard, D. (1990) *Empirical Processes: Theory and Applications*, Institute of Mathematical Statistics.

Primbs, J. (2009). Dynamic hedging of basket options under proportional transaction costs using receding horizon control. *International Journal of Control*, 82, 1841-1855.

Rachev, S.T., Menn, C., and Fabozzi, F.J. (2005). *Fat-Tailed and Skewed Asset Return Distributions: Implications for Risk Management, Portfolio Selection and Option Pricing*, WILEY.

Ramezani, C.A. and Zeng, Y. (1999). Maximum likelihood estimation of asymmetric jump-diffusion process: Application to security prices. *Working paper*.

Rogge, E., and Schnbucher, P.J. (2003). *Modelling dynamic portfolio credit risk*, Department of Mathematics, Imperial College and ABN AMRO Bank, London and Department of Mathematics, ETH Zurich, Zurich.

Schönbucher, P.J. (1998). Term structure modelling of defaultable bonds. *Review of Derivatives Research*, 2(2-3), 161-192.

Schönbucher, P.J. (2003). Information-driven default contagion. *Working Paper*, ETH, Zurich.

Schubert, D. and Schönbucher,, P.J. (2001). Copula-dependent default risk in intensity models. *Working paper*, Department of Statistics, Bonn University.

Shibata, T., and Yamada, T. (2009). Dynamic model of credit risk in relationship lending: a game-theoretic real options approach (No. 09-E-07). *Institute for Monetary and Economic Studies*, Bank of Japan.

Shiryaev, A. (1961). The detection of spontaneous effects. *Soviet Math. Dokl.*, 2, 740-743.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business*, 74(1), 101-124.

Soner, H.M., Shreve, S.E., and Cvitanić, J. (1995). There is no nontrivial hedging portfolio for option pricing with transaction costs. *Annals of Applied Probability*, 5, 327-355.

Tarashev, N.A. (2008). An empirical evaluation of structural credit-risk models. *International Journal of Central Banking*, 4(1), 1-53.

Vasicek, O. (1991). *Limiting loan loss probability distribution*, KMV corporation.

Vassalou, M., and Xing, Y. (2004). Default risk in equity returns. *The Journal of Finance*, 59(2), 831-868.

Whalley, A.E. and Wilmott, P. (1997). An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Mathematical Finance*, 7, 307-324.

Wilmott, P. (2009). *Frequently asked questions in quantitative finance*, WILEY.

Wilson, T. (1997a). Portfolio credit risk. I. *Risk*, 10(9):111-117

Wilson, T. (1997b). Portfolio credit risk. II. *Risk*, 10(10):56-61

Xing, H., Sun, N., and Chen, Y. (2012). Credit rating dynamics in the presence of unknown structural breaks. *Journal of Banking & Finance*, 36(1), 78-89.

Xing, H. and Yu, Y. (2013). Firms' default risk in the presence of economic structural breaks. *Working paper*

Zakamouline, V.I. (2006). European option pricing and hedging with both fixed and proportional transaction costs. *Journal of Economic Dynamics and Control*, 30(1), 1-25.