Jianwu Wang
Pricing Securities Subject to Credit Risk and Regime Switching
(Under the Direction of Qing Zhang)

This thesis begins with an introduction of credit risk and a review of credit risk models. A modified credit risk model which is subject to adjustable counterparty risks is then present, i.e. these counterparty risks are controlled by an exponential distribution. Following are two applications of this modified model on pricing credit derivatives: default swap of first-to-default baskets and collateralized bond obligations. Then a broader framework for the valuation of credit risk is introduced, i.e. incorporating Markov-modulated regime switching into the underlying factors of credit risk models. Under this generalized credit risk mode, two numerical methods: finite difference method and Markov Chain Monte Carlo simulation are used to calculate the prices of defaultable bonds. Finally, perpetual American put options subject to regime switching are studied. A stochastic approximation method is provided to find the optimal selling points for perpetual American put options.

Index words: American option, CBO, Copula function, Credit risk, Defaultable bond, Default swap, Finite difference method, Intensity process, Markov chain Monte Carlo simulation, Regime switching, Stochastic approximation approach.

Pricing Securities Subject to Credit Risk and Regime Switching

by

Jianwu Wang

B.S., Hangzhou Teachers' College, 1989

M.S., Sichuan University, 1992

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Athens, Georgia

2003

Pricing Securities Subject to Credit Risk and Regime Switching

by

Jianwu Wang

Approved:

Major Professor: Qing Zhang

Committee: Ralph E. Steuer
Ming-Jun Lai

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2003

Acknowledgments

I would like to thank my dissertation advisor Prof. Qing Zhang who introduced me to this interesting field of regime switching, helped me incorporate this regime switching into credit risk modelling. I really appreciate his constant encouragement, support and valuable advice.

I would like to thank my second committee member Prof. Ralph E. Steuer who offered me the chance to study at Terry College of Business (TCB), and monitored my graduate studies at TCB. He is always there when I need him. I'm very grateful for his kindness and support.

I would like to thank my third committee member Prof. Ming-Jun Lai for serving on my committees. I deeply appreciate his help.

I would also like to thank Prof. George Yin who introduced stochastic approximation approach to me. Both Prof. Yin and Prof. Zhang spent a lot time on explaining this approach to me. They are so generous to let me share their comments.

Finally, I dedicate this thesis to my wife, Hong Feng, who has provided huge support throughout my studies, and to my great son, Hengjia Wang.

Table of Contents

Chapter 1

General Introduction

1.1 Credit Risk and Credit Derivatives

Credit risk or *Default risk* is defined as "the risk that a counterparty defaults on its obligations". The government is generally assumed to meet the obligations of any financial contract it enters for certain. Any financial instruments issued by the government are therefore considered to be default free. However, financial contracts with counterparties other than the government are potentially default risky. Default risk is reduced by institutional arrangements in some markets. For example, organized securities exchanges have reduced default risk in futures contracts, options and other derivative securities by establishing clearing houses. However, such institutional arrangements for reducing default risk are not in place at the over-the-counter (OTC) markets or corporate bond markets. Default risk therefore affects corporate bonds as well as any securities traded at the OTC market.

Default risk is influenced by both business cycles and firm-specific events. Default risk typically declines during economic expansion because strong earnings keep overall defaults rates low. Default risk increases during economic recession because earnings deteriorate, making it more difficult to repay loans or make bond payments. Example here is the default of long-term capital management (LTCM), resulting from an adverse movement in interest rates. Default risk can also come from events specific to a firm's business activities, including the outcome of lawsuits, unexpected

devaluations, sudden default of a creditor, supplier, or a customer, and catastrophes in production lines. Barings is a case where a large trading loss forced bankruptcy. Therefore, default may be trigged by some unexpected events which cannot be observed from economic variables only.

Default risk affects the valuation and hedging of corporate bonds and all over-the-counter securities as well as portfolios of these securities. In 1992, a new class of securities called "Credit Derivatives" has been proposed at an annual meeting of the ISDA (International Swap Dealer Associations). *Credit derivatives* are securities "whose payoffs are linked to the credit characteristics of a particular asset". An example of a credit derivative is default swap, which pays a pre-specified amount in the event of default of a reference security in the swap contract. Now the credit derivatives market has grown dramatically. This growth has been driven by the ability of credit derivatives to provide valuable new methods for managing default risk. Credit derivatives can help banks, financial companies, and investors manage the default risk of their movements by insuring against adverse movements in the credit quality of the firm. If a firm defaults, the investor will suffer losses on the investment, but the losses can be offset by gains from the credit derivatives. Thus, in order to engineer credit derivative contracts to transfer the default risk exposure, the first thing is to develop a method to measure the default risk exposure correctly.

1.2 Review of Credit Risk Models

The pricing of defaultable bonds has been a major interest in the finance literature, and many models have been proposed for pricing and hedging risky debts. Among

them, there are two basic approaches for modelling default risks in bonds, structural models and reduced-form models.

1.2.1 STRUCTURAL FORM MODELS

The structural (or firm-value) approach was inspired by classic Black-Scholes option pricing theory. It assumes that the dynamics for the value of the assets of a firm across time can be described by a diffusion stochastic process and that the defaultable bond can be regarded as a contingent claim on the value of the assets of the firm. The structural approach is formulated by Merton [1974]. He assumed that the fundamental process V which represents the total value of the assets of the firm that has issued the bonds follows Geometric Brownian motion

$$\frac{dV}{V} = \mu dt + \sigma dW$$

A default occurs at maturity if V is insufficient to pay back the outstanding debt. Although Merton presented a breakthrough development in default risk pricing, there are many shortcomings of this model. The major shortcomings include: firm value is not observable; a flat and static yield spread; default occurring only when the firm value is less than the liability claim; default triggered only at the maturity of the debt; interest rate assumed constant over time.

There are some variations of the structural approach to overcome these shortcomings (see Longstaff and Schwartz [43], Zhou [60], etc.). In order to generate various shapes of yield spread curves, including upward-sloping, downward-sloping, flat, and hump-shaped, Zhou introduced jump into the underlying firm value process. Now the evolution of firm value follows a jump-diffusion process, so sudden drop in firm value becomes possible, therefore, default can occur unexpectedly. Longstaff and Schwartz [43] develop a more realistic model. They allow interest

rates to be stochastic. And they model default as the time when the value of the debt reaches some constant threshold value K that serves as a distress boundary. Contrary to Merton's model, default can occur prior to maturity. All these structural models have only limited success explaining the behavior of prices of debt instruments and credit spreads. These led to attempts to use models that make more direct assumptions on default process.

1.2.2 Reduced Form Models

Reduced-form models characterize default time as exogenously specified. The default time is unpredictable. Intensity function which is assumed to be determined by common economic factors as well as firm-specific factors is used to characterize default probability. The derived formulas of default time are calibrated to market data. This approach provides a model that is close to the date.

Lando [40] modelled the time of default as the first jump-time of a Cox process (also called doubly stochastic process). The random intensity of the Cox process may depend on interest rate or other factors.

Duffie and Singleton [13] developed a model where the payoff in default is assumed as a fraction of the value of the defaultable security just before default (called recovery at market value). Under their framework, defaultable security can proceed as in standard valuation models for default-free securities, using a default adjusted rate instead of the usual interest rate.

Other major papers include Hull & White [25] [26], Jarrow, Lando &Turnbull [28], Jarrow& Turnbull [29], Jarrow & Yu [31].

Chapter 2

Our Model Subject to counterparty Risk

2.1 Introduction

Jarrow & Yu [31] constructs a default intensity which can depend on firm-specific counterparty structures, in order to describe the default behavior of firms holding less well-diversified credit risk portfolios. They add a jump term in the intensity process when its counterparter suffers a default. In this chapter, we use copular function to derive that it's reasonable to adjust intensity process by adding a jump term if the two firms are correlated. Then we modify Jarrow & Yu [31] 's model to allow an exponential distribution to control this added term. So this added term may be dropped from the firm's intensity process in the future. It means that there is a possibility for the firm to get recovery after a certain period of time. Our model has the following features:

(i) As in Lando [40], Jeanblanc & Rutkowski [32], and Jarrow & Yu [31], we separate the information filtration into two parts: one filtration generated by state variables (interest rate and equity index) and the other filtration generated by the default process. So we can explicitly incorporate the correlation among the underlying firms into firms' default probability.

(ii) As Jarrow & Turnbull [30] pointed out, the issue of correlation is of central importance in all the credit risk methodologies. Two types of correlation are

often identified: default correlation and event correlation. Default correlation refers to firm default probabilities be correlated due to the common factors in the economy. Event correlation refers to how a firm's default probability is affected by default of other firms. We clearly discuss the default correlation and event correlation. We propose how to modify default probability when the event correlation is indispensable.

(iii) We use Farlie-Gumbel-Morgenstern copula to illustrate our results: one firm's default will help (deteriorate, resp.) its counterparty's survival if these firms are concordant (discordant, resp.). This result is the evidence to support the assumption in Jarrow & Yu [31] (they assume that there exists a jump term in the intensity function when its counterparty suffers a default). Meanwhile, I also find that one firm's survival will also help (deteriorate, resp.) its counterparty's survival if these firms are concordant (discordant, resp.).

(iv) Since the default event will affect its counterparty's intensity function, we don't assume that the default effect will be always on at its counterparty's rest time period. Instead, we assume that the holding time of this default effect follows exponential distribution. It means that the default effect will disappear after a certain time period.

(v) We let the event correlation be back into the intensity based credit risk model. This increases the the range of correlation in the model. So our model can remedy the disadvantage of the reduced-form model of limited range of correlation.

2.2 Classic Intensity-Based Models

We shall write $\mathbf{F}$ to denote a filtration $(\mathcal{F}_t,\ t\geqslant 0)$. In this paper, all filtrations are supposed to be completed and continuous on right, i.e. $\mathcal{F}_t = \cap_{t<s}\mathcal{F}_s$. Consider an economy indexed by the time interval [0,T^*]. Let the uncertainty in the economy be described by the filtered probability $(\Omega, \mathbf{F}, \mathbb{P})$. The probability space $(\Omega, \mathbf{F}, \{\mathcal{F}\}_{t=0}^{T^*}, \mathbb{P})$ is large enough to support a $\mathbf{R}^d$-valued stochastic process X, which we think of as the economic-wide factors (or state variables). A $\mathbf{R}_+ \cup \{+\infty\}$ valued random variable τ is an $\mathbf{G}$-stopping time if $\{\tau \leqslant t\} \in \mathcal{G}_t$, for any t, where $\mathbf{G} = (\mathcal{G}_t)_{\{t\geqslant 0\}}$. Obviously, if $\mathbf{F}$ is a filtration larger than $\mathbf{G}$, i.e. $\mathcal{G}_t \subset \mathcal{F}_t$ for any t, and τ is a $\mathbf{G}$-stopping time, then τ is a $\mathbf{F}$-stopping time, where $\mathbf{F}$ is a filtration of $(\mathcal{F}_t, t \geqslant 0)$. A stopping time τ is $\mathbf{F}$-*predictable* if there exists an increasing sequence of $\mathbf{F}$-stopping times τ_n such that $\tau_n < \tau$ on $\{\tau > 0\}$ and $\lim\tau_n = \tau$. A stopping time τ is $\mathbf{F}$-*totally inaccessible* if for any $\mathbf{F}$-predictable stopping time $\mathbf{S}, \mathbb{P}\{\omega \in \Omega : \tau(\omega) = \mathbf{S}(\omega) < \infty\} = 0$. In a Brownian filtration, it can be proved that any stopping time is a predictable stopping time. The most important example of totally inaccessible stopping time is the first time when a Poisson process jumps. In the intensity based credit risk model, we use Cox process (also doubly stochastic Poisson process) to represent default time. So the default time is totally inaccessible with respect to Brownian filtration. Otherwise, there'll be no intensity process for the default time. If τ is a nonnegative random variable on some probability space $(\Omega, \mathbf{F}, \mathbb{P})$, it is possible to endow Ω with a filtration such that τ is a stopping time. This filtration is not unique, and the right-continuous smallest filtration satisfying this property is $\mathcal{D}_t = \sigma(D_u, u \leqslant t)$, generated by the sets $\{\tau \leqslant s\}$ for s$\leqslant$t (that is the σ-algebra $\sigma(t \wedge \tau)$) and the atom $\{\tau > t\}$. Here $D_t = 1_{\{\tau\leqslant t\}}$ is the counting process associated with the random time τ. Notice that any $\mathcal{D}_t$-measurable integrable random variable H is of the form $\mathrm{H} = \mathrm{h}(\tau)1_{\{\tau\leqslant t\}} + \tilde{h}1_{\{\tau>t\}}$ where h is a Borel

function defined on [0, t], and $\tilde{h}$ a constant.

The filtration is generated collectively by the information contained in the state variables and the default process:

$\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{D}_t$ where $\mathcal{G}_t = \sigma(X_s, 0 \leqslant s \leqslant t)$ and $\mathcal{D}_t = \sigma(D_s, 0 \leqslant s \leqslant t)$ Where X_s is the state variable, i.e. interest rate, equity index, etc.

It is easy to describe the events which belong to the σ-field $\mathcal{F}_t$ on the set $\{\tau > t\}$. Indeed, if $A_t \in \mathcal{F}_t$, then $A_t \cap \{\tau > t\} = B_t \cap \{\tau > t\}$ for some event $B_t \in \mathcal{G}_t$. Therefore, any $\mathcal{F}_t$-measurable random variable Y_t satisfied $1_{\{\tau>t\}}Y_t = 1_{\{\tau>t\}}y_t$, where y_t is an $\mathcal{G}_t$-measurable random variable. In the filtrated probability space $(\Omega, \mathbf{F}, \{\mathcal{F}\}_{t=0}^{T^*}, \mathbb{P})$. Traded are default-free zero-coupon bonds of all maturities, a default-free money market account, and risky zero-coupon bonds of all maturities. We assume the market is complete and arbitrage free which means the defaultable claim is hedgeable. Under the assumption of no-arbitrage opportunities and complete, arbitrage pricing theory implies that there exists a unique equivalent martingale measure (e.m.m.) such that present value of a security is the expectation with respect to this e.m.m. discounted by the interest rate. In this paper, all calculation are under the e.m.m.

In the money market account accumulates returns at the spot rate and is denoted as $B(t) = \exp\left(\int_0^t r(s)ds\right)$ under the maintained assumption of arbitrage-free and complete markets, we can write default-free bond prices as the expected, defaultable value of a assure dollar received at time T, that is $p(t,T) = E\left(\frac{B(t)}{B(T)}\middle|\mathcal{F}_t\right)$ In the default risk framework, a default appears at some random time τ. The payment of a defaultable claims consists of two parts:

(1) Given a maturity date T > 0, a random variable Y, which does not dependent on τ represents the promised payoffs - that is, the amount of cash the owner of the claim will receive at time T, provided that the default has not occurred before the maturity date T.

(2) A predictable process h, prespecified in the default-free world, models the payoff which is received if default occurs before maturity. The process is called the recovery process or the rebate.

So the price of the defaultable claim is, provided that the default has not occurred before time t,

$$Y_t = E\left(Y1_{\{\tau>T\}}\exp\left(-\int_t^T r_u du\right) + h_\tau 1_{\{\tau\leqslant T\}}\exp\left(-\int_t^T r_u du\right)\middle|\mathcal{F}_t\right)$$

where $\mathcal{F}_t$ is all the information up to time t, r_u is the spot interest rate. Under the existence of intensity of the default, we'll see that the intensity of the default time acts as a change of the spot interest rate in the pricing formula.

For any $t \in \mathbb{R}_+$ and firm i, we denote $F_t^i = \mathbb{P}(\tau_i \leqslant t|\mathcal{G}_t)$ the conditional default probability of firm i given the state variables . Assume that $F_t^i < 1$ for every $t \in \mathbb{R}_+$. The $\mathcal{G}$-*hazard process* of τ_i, denoted by Γ^i, is defined by the formula $1 - F_t^i = e^{-\Gamma_t^i}$, so the *conditional survival probability* $\mathbb{P}(\tau_i \geqslant t|\mathcal{G}_t)$ is equal to $e^{-\Gamma_t^i}$, and Γ_t^i is $\mathcal{G}_t$-measurable. In this paper, we assume that the cumulative distribution function F_t^i is absolutely continuous, that is, $F_t^i = \int_0^t f_i(s)ds$, for some function $f_i : \mathbb{R}_+ \to \mathbb{R}_+$. So we have

$$F_t^i = 1 - e^{-\Gamma_t^i} = 1 - e^{-\int_0^t \lambda_i(s)ds},$$

where *intensity process* $\lambda_i(t) = \frac{f_i(t)}{1-F_t^i}$, and $\Gamma_t^i = \int_0^t \lambda_i(s)ds$. The interpretation of the intensity process is that over the interval (t, t+Δt] the default probability conditional

upon no default prior to time t is approximately $\lambda_i(t)\Delta$t. We define the default time τ_i as follows

$$\tau_i = \inf\{t : e^{-\Gamma_t^i} \leqslant E_i\}$$

where E_i is uniform random variable which is assumed to be independent with $\mathcal{G}_{T^*}$. As in Lando [40], this default time can be thought of as the first jump time of a Cox process with intensity process $\lambda_i(s)$. It is obvious that we have $\{\tau_i \leqslant t\} = \{E_i \geqslant e^{-\Gamma_t^i}\}$, $\{\tau_i > t\} = \{E_i < e^{-\Gamma_t^i}\}$ and

$$\mathbb{P}(\tau_i > t|\mathcal{G}_t) = \exp(-\int_0^t \lambda_i(s)ds)$$

And for $u \geqslant t$

$$\mathbb{P}(\tau_i > u|\mathcal{G}_t) = E(E(1_{\{\tau_i>u\}}|\mathcal{G}_{T^*})|\mathcal{G}_t) = E(\exp(-\int_0^u \lambda_i(s)ds)|\mathcal{G}_t)$$

Note:(i) The intensity process $\lambda_i(t)$ is actually a function of state variable X_t. We drop the symbol X for convenience.

(ii) The independency between E_i and $\mathcal{G}_{T^*}$ implies that τ_i is not measurable with respect to $\mathcal{G}_t$, so τ_i is not **G**-stopping time, therefore $F_t^i \stackrel{def}{=} \mathbb{P}(\tau_i \leqslant t|\mathcal{G}_t)$ is well-defined.

Check the appendix to see the proof of the following equations:

(a) For any $\mathcal{F}$-measurable random variable Y, we have, for any $t \in R_t$

$$E(1_{\{\tau_i>s\}}Y|\mathcal{F}_t) = 1_{\{\tau_i>t\}}\frac{E(1_{\{\tau_i>s\}}Y|\mathcal{G}_t)}{\mathbb{P}(\tau_i > t|\mathcal{G}_t)} \tag{2.1}$$

for any $t \leqslant s$, In particular, we have $\mathbb{P}(\tau_i > s|\mathcal{F}_t) = 1_{\{\tau_i>t\}}E(\exp(-\int_t^s \lambda_i(u)du)|\mathcal{G}_t)$.

(b) Let $h : R_+ \to R$ be a (bounded) Borel measurable function. Then

$$E(1_{t<\tau_i\leqslant s}h(\tau_i)\big|\mathcal{G}_t) = 1_{\{\tau_i>t\}}\frac{E\left(\int_{]t,s]} h(u)dF_u\Big|\mathcal{G}_t\right)}{\mathbb{P}(\tau_i > t\big|\mathcal{G}_t)} \tag{2.2}$$

where $F_u = \mathbb{P}(\tau_i \leqslant u \big| \mathcal{G}_t)$

(c) In the general case, let Z be a (bounded) **G**-predictable process, then for any $t \leqslant s$

$$E(1_{\{t<\tau_i\leqslant s\}} Z_{\tau_i} | \mathcal{F}_t) = 1_{\{\tau_i > t\}} \frac{E(\int_{]t,s]} Z_u dF_u | \mathcal{G}_t)}{\mathbb{P}(\tau_i > t | \mathcal{F}_t)} \tag{2.3}$$

From the above lemma, we have following equations (which also can be found in Lando [40]) :

$$\begin{aligned}
&E\left(X 1_{\{\tau_i>T\}} \exp\left(-\int_t^T r(s)ds\right) \Big| \mathcal{F}_t\right) \\
&= 1_{\{\tau_i>t\}} E\left(X \exp\left(-\int_t^T (r(s)+\lambda_i(s))ds\right) \Big| \mathcal{G}_t\right) \\
&E\left(\int_t^T Y_s 1_{\{\tau_i>s\}} \exp\left(-\int_t^s r(u)du\right) ds \Big| \mathcal{F}_t\right) \\
&= 1_{\{\tau_i>t\}} E\left(\int_t^T Y_s \exp\left(-\int_t^s (r(u)+\lambda_i(u))du\right) ds \Big| \mathcal{G}_t\right) \\
&E\left(\exp\left(-\int_t^{\tau_i} r(s)ds\right) Z_{\tau_i} \Big| \mathcal{F}_t\right) \\
&= 1_{\{\tau_i>t\}} E\left(\int_t^T Z_s \lambda_i(s) \exp\left(-\int_t^s (r(u)+\lambda_i(u))du\right) ds \Big| \mathcal{G}_t\right).
\end{aligned} \tag{2.4}$$

Note (i) The interpretation of the above equations is that a rational investor should be indifferent between the expected cash flows discounted by risk free rate (i.e. left terms of these equations) and the promised cash flows discounted by the risky rate. (ii) As we pointed out before, the intensity of the default time acts as a change of

the spot interest rate in the pricing formula.

Suppose bond i's recovery rate is δ_i, and we use the recovery of Treasury assumption proposed in Jarrow & Turnbull [29]. Using the above equations, the time t price of defaultable zero-coupon bond is

$$v_i(t,T) = E\left(1_{\{t<\tau_i\leqslant T\}}\delta_i \exp\left(-\int_t^T r(s)ds\right) + 1_{\{\tau_i>T\}}\exp\left(-\int_t^T r(s)ds\right)\Big|\mathcal{F}_t\right)$$

$$= \delta_i p(t,T) + (1-\delta_i)1_{\{\tau_i>t\}}E\left(\exp\left(-\int_t^T [r(s)+\lambda_i(s)]ds\right)Y\Big|\mathcal{G}_t\right) \tag{2.5}$$

If the intensity function $\lambda_i(t)$ is a constant, say λ_0, then the above formula can be simplified as follow:

$$v_i(t,T) = \delta_i p(t,T) + (1-\delta_i)1_{\{\tau_i>t\}}p(t,T)e^{-\lambda_0(T-t)} \tag{2.6}$$

Or

$$\frac{v_i(t,T)}{p(t,T)} = \delta_i + (1-\delta_i)1_{\{\tau_i>t\}}e^{-\lambda_0(T-t)} \tag{2.7}$$

This kind of credit risk model which fits for one individual firm was studied extensively by Lando [40].

2.3 Copula Functions

In order to introduce correlation structure into the n risky bonds, there is no unique solution. We choose copula function to describe the joint distribution of the n dependent default times.

2.3.1 Definition and Notation

Copula is a multivariate distribution function defined on the unit cube $[0,1]^n$, with uniformly distributed marginal. Let $S_1, \cdots, S_n$ be nonempty subsets of $\bar{\mathbf{R}}$, where $\bar{\mathbf{R}}$

denotes the extended real line $[-\infty, \infty]$. Let H be the a real function of n variables such that $DomH = S_1 \times \cdots \times S_n$ and let $B = [a, b]$ be an n-box which vertices are in $DomH$. Then the *H-volume of B* is given by

$$V_H(B) = \sum sgn(\mathbf{c})H(\mathbf{c}),$$

where the sum is taken over all vertices $\mathbf{c}$ of B, and $sgn(\mathbf{c})$ is given by

$$sgn(\mathbf{c}) = \begin{cases} 1, & \text{if } c_k = a_k \text{ for an even number of k's,} \\ -1, & \text{if } c_k = a_k \text{ for an odd number of k's.} \end{cases}$$

Equivalently, the H-volume of an n-box $B = [\mathbf{a}, \mathbf{b}]$ is the n^{th} order difference of H on B

$$V_H(B) = \Delta_{\mathbf{a}}^{\mathbf{b}} H(\mathbf{t}) = \Delta_{a_n}^{b_n} \cdots \Delta_{a_1}^{b_1} H(\mathbf{t})$$

where we define the n first order difference as

$$\Delta_{a_k}^{b_k} H(\mathbf{t}) = H(t_1, \ldots, t_{k-1}, b_k, t_{k+1}, \ldots, t_n) - H(t_1, \ldots, t_{k-1}, a_k, t_{k+1}, \ldots, t_n)$$

A real function H of n variables is *n-increasing* if $V_H(B) \geqslant 0$ for all n-boxes B whose vertices lie in $DomH$.

Consider $n = 2$, we have

$$V_H(B) = \Delta_{y_1}^{y_2} \Delta_{x_1}^{x_2} H(x, y) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1)$$

and

$$\Delta_{x_1}^{x_2} H(x, y) = H(x_2, y) - H(x_1, y) \qquad \text{and} \qquad \Delta_{y_1}^{y_2} H(x, y) = H(x, y_2) - H(x, y_1)$$

Note that the statement "H is 2-increasing" neither implies nor is implied by the statement "H is nondecreasing in each argument," as the following two examples illustrate.

Example 1 *Let H be the function defined on* $[0, 1]^2$ *by* $H(x, y) = \max(x, y)$. *Then H is a nondecreasing function of x and of y; however,* $V_{[0,1]^2} = -1$, *so that H is not 2-increasing.*

Example 2 *Let H be the function defined on* $[0,1]^2$ *by H(x,y)= (2x-1)(2y-1). Then H is 2-increasing, however it is a decreasing function of x for each y in (0,1/2) and a decreasing function of y for each x in (0,1/2).*

Suppose that the domain of a real function H of n variables is given by $DomH = S_1 \times \cdots \times S_n$ where each S_k has a smallest element a_k. We say that H is *grounded* if $H(\mathbf{t}) = 0$ for all $\mathbf{t}$ in $DomH$ such that $t_k = a_k$ for at least one k. If each S_k is nonempty and has a greatest element b_k, then H has *margins*, and the one-dimensional margins of H are the functions H_k with $DomH_k = S_k$, and for all x in S_k, $H_k(x) = H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n)$. Higher dimensional margins are defined in an obvious way.

An *n-dimensional copula* is a function C with domain $[0,1]^n$ such that
(i). C is grounded and n-increasing.
(ii). C has margins $C_k, k = 1, 2, \ldots, n$, which satisfy $C_k(u) = u$ for all u in [0,1].
Note that for any n-copula C, $n \geqslant 3$, each k-dimensional of C is a k-copula. Equivalently, an n-copula is a function C from $[0,1]^n$ to [0,1] with the following properties:
(a). For every $\mathbf{u}$ in $[0,1]^n$, C($\mathbf{n}$) = 0 if at least one coordinate of $\mathbf{u}$ is 0, and C($\mathbf{u}$) $= u_k$ if all coordinates of $\mathbf{u}$ equal 1 except u_k.
(b). For every $\mathbf{a}$ and $\mathbf{b}$ in $[0,1]^n$ such that $a_i \leqslant b_i$ for all i, $V_C(\mathbf{a}, \mathbf{b}) \geqslant 0$.

The following theorem is known as Sklar's theorem. It is the most important result regarding copulas, and is used in essentially all applications of copulas.

Sklar's theorem [Nelsen [48]] Let $\mathbb{F}$ be an n-dimensional distribution function with margins $F_1, \ldots, F_n$. Then there exists an n-copula C such that for all $\mathbf{x}$ in $\bar{\mathbf{R}}^n$,

$$\mathbb{F}(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)).$$

If $F_1, \ldots, F_n$ are all continuous, then C is unique; otherwise C is uniquely determined on $RanF_1 \times \cdots \times RanF_n$. Conversely, if C is an n-copula and $F_1, \ldots, F_n$ are distribution functions, then the function $\mathbb{F}$ defined above is an n-dimensional function with margins $F_1, \ldots, F_n$.

From Sklar's theorem we see that for continuous multivariate distribution functions, the univariate margins and the multivariate structure can be separated, and the dependence structure can be represented by a copula.

One nice property of copula is that for strictly monotone transformation of the random variables, copula is either invariant, or change in certain simple way. It means that if $(X_1, \ldots, X_n)$ be a vector of continuous random variables with copula C and $\beta_1, \ldots, \beta_n$ are strictly increasing on $RanX_1, \ldots, RanX_n$, respectively, then also $(\beta_1(X_1), \ldots, \beta_n(X_n))$ has copula C. Copula provides a natural way to study and measure dependence between random variables. Both Spearman's Rho and Kendall's Tau can be use to indicate the correlation between two random variables. Spearman's Rho and Kendall's Tao can be defined by using copula function. We give the formula here.

$$Rho = 12 \iint_{[0,1]^2} uvdC(u,v) - 3 = 12 \iint_{[0,1]^2} C(u,v)dudv - 3.$$

and

$$Tau = 4 \iint_{[0,1]^2} C(u,v)dC(u,v) - 1.$$

2.3.2 Linear Correlation vs Copula

Copula provides a natural way to study and measure dependence between random variables. As we know, linear correlation[1] (or Pearson's correlation) is also fre-

[1]Let X and Y be two random variables with finite variances. The linear correlation coefficient for X and Y is $\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$

quently used in practice as a measure of dependence. The popularity of linear correlation stems from the ease with which it can be calculated and it is a natural scalar measure of dependence in elliptical distributions (The elliptical distributions are distributions whose density is constant on ellipsoids. In two dimensions, the contour lines of the density surface are ellipses. The multivariate normal is a special case). However, most random variables are not jointly elliptically distributed, and using linear correlation as a measure of dependence in such situations might prove very misleading. Even for jointly elliptically distributed random variables there are situations where using linear correlation does not make sense. We might choose to model some scenario using heavy-tailed distributions such as t_2-distributions. Correlation tells us nothing about the degree of dependence in the tail of the underlying distribution. In such cases the linear correlation coefficient is not even defined because of infinite second moments.

A list of the problems of linear correlation as a dependency measure is:

(a) Linear correlation is simple a scalar measure of dependence. It cannot tell us everything we would like to know about the dependence structure of risks.

(b) Possible values of linear correlation depend on the marginal distribution of the risks. All values between -1 and 1 are not necessarily attainable.

(c) Perfectly positively dependent risks do not necessarily have a correlation of 1. Perfectly negatively dependent risks do not necessarily have a correlation of -1.

(d) A linear correlation of zero does not indicate independence of risks.

(e) Linear correlation is not invariant under transformations of risks. For example, $\log(X)$ and $\log(Y)$ generally do not have the same linear correlation as X and Y.

(f) Linear correlation is only defined when the variances of the risks are finite. It is not an appropriate dependence measure for very heavy-tailed where variances appear infinite.

By turning to rank correlation, certain of these theoretical deficiencies of standard linear correlation can be repaired. It does not matter whether we choose the Kendall's Tau or Spearman's Rho definitions of rank correlation. Rank correlation does not have deficiencies (b), (c), (e), and (f). Copulas represent a way of trying to extract the dependence structure from the joint distribution and to extricate dependence and marginal behavior.

2.3.3 Concordance and Discordance

Informally, a pair of random variables are concordant if "large" values of one tend to be associated with "large" values of the other, and "small" values of one with "small" values of the other. To be more precise, let (x_i, y_i) and (x_j, y_j) denote two observations from a vector (X, Y) of continuous random variables. We say that (x_i, y_i) and (x_j, y_j) are *concordant* if $x_i < x_j$ and $y_i < y_j$, or if $x_i > x_j$ and $y_i > y_j$. Similarly, we say that (x_i, y_i) and (x_j, y_j) are *discordant* if $x_i < x_j$ and $y_i > y_j$, or if $x_i > x_j$ and $y_i < y_j$. Note the alternate formulation: (x_i, y_i) and (x_j, y_j) are concordant if $(x_i - x_j)(y_i - y_j) > 0$, and discordant if $(x_i - x_j)(y_i - y_j) > 0$. We use Kendall's tau or Spearman's rho to describe concordant and discordant.

Kendall's tau for the random variables X and Y is defined as

$$Tau(X,Y) = \mathbb{P}\{(X-\tilde{X})(Y-\tilde{Y})>0\} - \mathbb{P}\{(X-\tilde{X})(Y-\tilde{Y})<0\}$$

where $(\tilde{X},\tilde{Y})$ is an independent copy of (X,Y). This is the population version of Kendall's tau which is defined as the probability of concordance minus the probability of discordance. However, Kendall's Tau can also be defined using a copula function only. From Nelsen [48], we know

$$Tau(X,Y) = 4\iint_{[0,1]^2} C(u,v)dC(u,v) - 1.$$

Similarly, we give the population version and copula version of Spearman's Rho's definition just for reference. In practice, we can choose either Kendall's Tao or Spearman's Rho. Spearman's rho for the random variables X and Y is defined as

$$Rho(X,Y) = 3(\mathbb{P}\{(X-\tilde{X})(Y-Y')>0\} - \mathbb{P}\{(X-\tilde{X})(Y-Y')<0\})$$

where (X,Y),$(\tilde{X},\tilde{Y})$ and (X',Y') are independent copies. And the copula version is:

$$Rho(X,Y) = 12\iint_{[0,1]^2} uvdC(u,v) - 3 = 12\iint_{[0,1]^2} C(u,v)dudv - 3.$$

For Farlie-Gumbel-Morgenstern copula $C(u,v) = uv(1+\alpha(1-u)(1-v)), \alpha\in[-1,1]$, we know the two random variables are independent when $\alpha = 0$. And Kendall's $Tau = \frac{2}{9}\alpha$, Spearmen's $Rho = \frac{1}{3}\alpha$. From these simple forms of Kendall's Tau and Spearmen's Rho, we know that the two random variables will move in the same direction when $\alpha > 0$, and will move in the opposite direction when $\alpha < 0$.

2.4 Dependent Default Models

In section 2.2, we've derived the default risk model for individual firm. Now we'll derive the default risk model for the case when we consider n firms simultaneously. We denote the counting process associated with the default time τ_i of firm i by

$D_t^i = 1_{\{\tau_i \leqslant t\}}$ and its σ-field by $\mathcal{D}_t^i$ ($\stackrel{def}{=} \sigma(D_u^i, u \leqslant t)$). Now all the information up to time which is available to investors is

$$\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{D}_t^1 \vee \cdots \vee \mathcal{D}_t^n$$

As in section 2.2, default time τ_i is defined as follows:

$$\tau_i = \inf\{t : e^{-\Gamma_t^i} \leqslant E_i\}$$

where E_i is uniform random variable which is assumed to be independent with $\mathcal{G}_{T^*}$. Now we have n default times $(\tau_1, \tau_2, \cdots, \tau_n)$. In Kijima [35], he assumes that these n default times are conditional independent, i.e.

$$\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2, \cdots, \tau_n > t_n | \mathcal{G}_{T^*}) = \prod_{i=1}^{n} \mathbb{P}(\tau_i > t_i | \mathcal{G}_{T^*})$$

Since we know that $\mathbb{P}(\tau_i > t_i | \mathcal{G}_{t_i}) = \exp(-\int_0^{t_i} \lambda_i(s)ds)$. So using intensity processes, the conditional independent assumption will give the following equation:

$$\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2, \cdots, \tau_n > t_n | \mathcal{G}_{T^*}) = \exp(-\sum_{i=1}^{n} \int_0^{t_i} \lambda_i(s)ds)$$

This assumption will simplify the question a lot when we deal with the joint distribution of these default times. But, we'll see that this assumption will sacrifice a lot important information. So we don't assume that the n default times $\tau_1, \tau_2, \cdots, \tau_n$ are conditional independent. According to the construction of default times, we have the following equation:

$$\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2, \cdots, \tau_n > t_n) = \mathbb{P}(E_1 < e^{-\Gamma_{t_1}^1}, E_2 < e^{-\Gamma_{t_2}^2}, \cdots, E_n < e^{-\Gamma_{t_n}^n})$$

where the hazard process $\Gamma_{t_i}^i = \int_o^{t_i} \lambda_i(s)ds$ and E_i is uniform random variable. From the above equation, we know that describing the joint distribution of random variables E_i is equivalent to describing the joint distribution of default times τ_i. Now

we suppose that the copula function of these n random variables E_i is C. So we have the following:

$$\begin{aligned}\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2, \cdots, \tau_n > t_n) &= \mathbb{P}(E_1 < e^{-\Gamma^1_{t_1}}, E_2 < e^{-\Gamma^2_{t_2}}, \cdots, E_n < e^{-\Gamma^n_{t_n}}) \\ &= C(e^{-\Gamma^1_{t_1}}, e^{-\Gamma^2_{t_2}}, \cdots, e^{-\Gamma^n_{t_n}})\end{aligned}$$

2.4.1 Default Correlation

Default correlation (see, Jarrow & Turnbull [30]) refers to firm default probabilities being correlated due to common factors in the economy. We use interest rate $r(t)$ as the common factor to see how it will affect the default correlation even when there is no correlation between the firm (i.e. $\alpha = 0$) . let the their intensity functions $\lambda_A(t)$ and $\lambda_B(t)$ be functions of spot rate of interest:

$$\lambda_A(t) = \lambda_0^A + \lambda_1^A r(t) \quad \text{and} \quad \lambda_B(t) = \lambda_0^B + \lambda_1^B r(t)$$

where $\lambda_0^A, \lambda_1^A, \lambda_0^B, \lambda_1^B$ are constants. so their survival probabilities are:

$$\mathbb{P}(\tau_A > t|\mathcal{G}_t) = e^{-\int_0^t \lambda_A(s)ds} = e^{-\int_0^t (\lambda_0^A + \lambda_1^A r(s))ds}$$

And

$$\mathbb{P}(\tau_B > t|\mathcal{G}_t) = e^{-\int_0^t \lambda_B(s)ds} = e^{-\int_0^t (\lambda_0^B + \lambda_1^B r(s))ds}$$

As in Jarrow & Turnbull [30], we assume the spot rate of interest follow the extended Vasicek model:

$$dr(t) = a(\overline{r}(t) - r(t))dt + \sigma_r dW_r(t)$$

where $W_r(t)$ is a Wiener process under the e.m.m. $\mathbb{P}$, and $\overline{r}(t)$ is a deterministic function chosen to fit an initial term structure, with a and σ_r constants. From Jarrow & Turnbull [30], Jarrow [27], we have:

$$\overline{r}(t) = f(0,t) + \frac{1}{a}\frac{\partial f(0,t)}{\partial t} + \frac{\sigma_r^2}{2a^2}(1 - e^{-2at})$$

And

$$\mu_{t,T} = E(\int_t^T r(u)du|\mathcal{G}_t) = \int_t^T f(t,u)du + \int_t^T \frac{b(u,T)^2}{2}du$$

$$\sigma_{t,T}^2 = \text{var}(\int_t^T r(u)du|\mathcal{G}_t) = \int_t^T b(u,T)^2 du$$

where $f(t,u)$ is the forward rate, and $b(u,T) = \frac{\sigma_r}{a}(1-e^{-a(T-u)})$. Therefore, we have:

$$\mathbb{P}(\tau_A > t) = E[\mathbb{P}(\tau_A > t|\mathcal{G}_t)] = E[e^{-\int_0^t \lambda_A(s)ds}] = e^{-\lambda_0^A t - \lambda_1^A \mu_{0,t} + \frac{(\lambda_1^A)^2}{2}\sigma_{0,t}^2}$$

And

$$\mathbb{P}(\tau_A > t) = E[\mathbb{P}(\tau_B > t|\mathcal{G}_t)] = E[e^{-\int_0^t \lambda_B(s)ds}] = e^{-\lambda_0^B t - \lambda_1^B \mu_{0,t} + \frac{(\lambda_1^B)^2}{2}\sigma_{0,t}^2}$$

Now in order to see the default correlation, we let $\alpha = 0$. So

$$\mathbb{P}(\tau_A > t, \tau_B > t) = E[\mathbb{P}(\tau_A > t, \tau_B > t|\mathcal{G}_{T^*})]$$

$$= E[e^{-\int_0^t \lambda_A(u)du} e^{-\int_0^t \lambda_B(u)du}]$$

$$= E[e^{-\int_0^t (\lambda_0^A + \lambda_1^A r(s))ds} e^{-\int_0^t (\lambda_0^B + \lambda_1^B r(s))ds}]$$

$$= e^{-(\lambda_0^A+\lambda_1^B)t - (\lambda_1^A+\lambda_1^B)\mu_{0,t} + \frac{(\lambda_1^A+\lambda_1^B)^2}{2}\sigma_{0,t}^2}$$

Comparing the above equations, we have:

$$\mathbb{P}(\tau_A > t, \tau_B > t) = \mathbb{P}(\tau_A > t)\mathbb{P}(\tau_B > t) e^{\lambda_1^A \lambda_1^B \sigma_{0,t}^2}$$

From the above equation, we know that if none of λ_1^A and λ_1^B is zero (it means that both of these two bonds depends on the common factor–interest rate), we do have default correlation because of the extra term $e^{\lambda_1^A \lambda_1^B \sigma_{0,t}^2}$.

2.4.2 Effect of Event Correlation

Event correlation refers to how a firm's default probability is affected by default of other firms (see Jarrow & Turnbull [30]). Now, let's check how one firm's default will affect its counterparty's default. For simplicity, we consider two bonds issued by firm A and firm B. We use τ_A, τ_B to denote their default times, $\lambda_A(t), \lambda_B(t)$ to denote their intensity processes (so hazard processes are $\Gamma_t^A = \int_0^t \lambda_A(u)du, \Gamma_t^B = \int_0^t \lambda_B(u)du$), and $\mathcal{D}_t^A, \mathcal{D}_t^B$ to denote their information filtrations generated by default times τ_A, τ_B, resp. We know that if these two firms are independent (i.e. random variable τ_B is independent of $\mathcal{D}_t^A$), then we have $\mathbb{P}(\tau_B > s|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = \mathbb{P}(\tau_B > s|\mathcal{G}_t \vee \mathcal{D}_t^B)$, for $s \geqslant t$. It means that the information of $\mathcal{D}_t^A$ has no effect on the survival probability of bond B.

Now we assume that the two firms are correlated [2], the information which is available to firm B at time t is $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B$, and the information of τ_A (i.e. $\mathcal{D}_t^A$) will affect the default probability of bond B. Now let's use copula function to see the effect of event correlation. Since Farlie-Gumbel-Morgenstern copula has a very simple formula of Kendall's Rho and Spearman's Rho, we can easily determine the two firms' dependency based on the parameter α, i.e. if the parameter α is positive, then the two random variables are concordant, otherwise the two random variables are discordant.

[2] The common factors serve to induce the default correlation between firms. However, this default correlation is very small if the firms are from different industry sectors. That's why Moody's treats firms from different industry sectors as independent. But when two firms are from same industry sectors, it's not reasonable to assume they are independent.

The appendix gives the proof of the following equation (for $s \geqslant t$):

$$\mathbb{P}(\tau_B > s \big| \mathcal{G}_t^B) = \mathbb{P}(\tau_B > s \big| \mathcal{G}_t \vee \mathcal{D}_t^A)$$

$$= \begin{cases} e^{\Gamma_t^A} E[C(e^{-\Gamma_s^B}, e^{-\Gamma_t^A}) \big| \mathcal{G}_t], & if \quad 1_{\{\tau_A > t\}} = 1 \\ \frac{E\left(e^{-\Gamma_s^B} \big| \mathcal{G}_t\right) - E[C(e^{-\Gamma_s^B}, e^{-\Gamma_t^A}) \big| \mathcal{G}_t]}{1 - e^{-\Gamma_t^A}}, & if \quad 1_{\{\tau_A \leqslant t\}} = 1. \end{cases} \tag{2.8}$$

Now let's specify the copula in the above equation by Farlie-Gumbel-Morgenstern copula, so we have:

$$\mathbb{P}(\tau_B > s \big| \mathcal{G}_t \vee \mathcal{D}_t^A)$$

$$= \begin{cases} E[e^{-\Gamma_s^B}(1 + \alpha(1 - e^{-\Gamma_s^B})(1 - e^{-\Gamma_t^A})) \big| \mathcal{G}_t], & if \quad 1_{\{\tau_A > t\}} = 1 \\ E[e^{-\Gamma_s^B}\left(1 - \alpha e^{-\Gamma_t^A}(1 - e^{-\Gamma_s^B})\right) \big| \mathcal{G}_t], & if \quad 1_{\{\tau_A \leqslant t\}} = 1. \end{cases} \tag{2.9}$$

From the above equation, we have the following results:

(1) When $\alpha = 0$, either case reduces to $E\left(e^{-\Gamma_s^{AB}} \big| \mathcal{G}_t\right)$ which is equal to $\mathbb{P}(\tau_B > s \big| \mathcal{G}_t)$. This goes back to the case where the firms are independent.

(2) If $1_{\{\tau_A > t\}} = 1$ and $\alpha > 0$, then $1 + \alpha(1 - e^{-\Gamma_s^B})(1 - e^{-\Gamma_t^A}) > 1$. So we have $E[e^{-\Gamma_s^B}(1 + \alpha(1 - e^{-\Gamma_s^B})(1 - e^{-\Gamma_t^A})) \big| \mathcal{G}_t] > E\left(e^{-\Gamma_s^B} \big| \mathcal{G}_t\right) (= \mathbb{P}(\tau_B > s \big| \mathcal{G}_t))$. It means that the survival of firm A will increase the survival probability of firm B if firm A and firm B are concordant ($\alpha > 0 \Rightarrow Tau = \frac{2}{9}\alpha > 0$). For example, since Intel is the main supplier of key part of computer for PC manufacture, say Compaq, so Intel and Compaq are concordant. Our result shows that the survival of Intel will help Compaq to survive.

(3) If $1_{\{\tau_A>t\}} = 1$ and $\alpha < 0$, then $1 + \alpha(1 - e^{-\Gamma_s^B})(1 - e^{-\Gamma_t^A}) < 1$. So we have $E[e^{-\Gamma_s^B}(1 + \alpha(1 - e^{-\Gamma_s^B})(1 - e^{-\Gamma_t^A}))\big|\mathcal{G}_t] < E\left(e^{-\Gamma_s^B}\big|\mathcal{G}_t\right) (= \mathbb{P}(\tau_B > s\big|\mathcal{G}_t))$. It means that the survival of firm A will decrease the survival probability of firm B if firm A and firm B are discordant ($\alpha < 0 \Rightarrow Tau = \frac{2}{9}\alpha < 0$). For example, telecom companies AT&T and Sprint are competitors. Their relationship can be considered discordant. So our result says that the survival of AT&T will deteriorate the survival of Sprint.

(4) If $1_{\{\tau_A \leqslant t\}} = 1$ and $\alpha > 0$, then $1 - \alpha e^{-\Gamma_t^A}(1 - e^{-\Gamma_s^B}) < 1$. So we have

$$E\left(e^{-\Gamma_s^B}\left(1 - \alpha e^{-\Gamma_t^A}(1 - e^{-\Gamma_s^B})\right)\Big|\mathcal{G}_t\right) < E\left(e^{-\Gamma_s^B}\big|\mathcal{G}_t\right)$$

$(= \mathbb{P}(\tau_B > s\big|\mathcal{G}_t))$. It means that the default of firm A will reduce the survival probability of firm B if they are concordant. Using the above example, our result shows that the default of Intel will cause the the survival probability of Compaq to decrease.

(5) If $1_{\{\tau_A \leqslant t\}} = 1$ and $\alpha < 0$, then $1 - \alpha e^{-\Gamma_t^A}(1 - e^{-\Gamma_s^B}) > 1$. So we have

$$E\left(e^{-\Gamma_s^B}\left(1 - \alpha e^{-\Gamma_t^A}(1 - e^{-\Gamma_s^B})\right)\Big|\mathcal{G}_t\right) > E\left(e^{-\Gamma_s^B}\big|\mathcal{G}_t\right)$$

$(= \mathbb{P}(\tau_B > s\big|\mathcal{G}_t))$. It means that the default of firm A will increase the survival probability of firm B when they are discordant. Using the above example, our result Sprint shows that the default of AT&T will help to survive.

From the above results, we find that one firm's information (default or survival) does affect its counterparty's default probability when they are correlated (i.e. $\alpha \neq 0$), even when there is no default event occurs. So we extend the event correlation to this general sense.

Now let's use the approximation $e^x \doteq 1 + x$ to see why Jarrow & Yu [31]'s assumption is reasonable, i.e. add one term to the firm's intensity process when its counterparty gets default. For $s \geqslant t$, and check the Jeanblanc & Rutkowski [32] , we have:

$$\mathbb{P}(\tau_B > s|\mathcal{G}_t \vee \mathcal{D}_t^B) = 1_{\{\tau_A > t\}} \frac{\mathbb{P}(\tau_B > s|\mathcal{G}_t)}{\mathbb{P}(\tau_b > t|\mathcal{G}_t)} = E(e^{-\int_t^s \lambda_B(u)du}|\mathcal{G}_t) \tag{2.10}$$

And

$$\begin{aligned}\mathbb{P}(\tau_B > s|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) &= 1_{\{\tau_B > t, \tau_A \leqslant t\}} \frac{\mathbb{P}(\tau_B > s, \tau_A \leqslant t|\mathcal{G}_t)}{\mathbb{P}(\tau_B > t, \tau_A \leqslant t|\mathcal{G}_t)} \\ &= 1_{\{\tau_B > t, \tau_A \leqslant t\}} E(e^{-\int_t^s \lambda_B(u)(1+\alpha e^{-\int_0^t \lambda_A(v)dv})du}|\mathcal{G}_t)\end{aligned} \tag{2.11}$$

Comparing the above two equations, we see that there does have a term $\alpha e^{-\int_0^t \lambda_A(v)dv}$ added into bond B's intensity function. How this term affects the bond B's survival probability depends on the sign of α.

Similarly, appendix gives the derivation of the following equation under the approximation of $e^x \doteq 1 + x$, for the case when its counterparty has survived up to time t.

$$\begin{aligned}\mathbb{P}(\tau_B > s|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) &= 1_{\{\tau_A > t, \tau_B > t\}} \frac{\mathbb{P}(\tau_A > t, \tau_B > s|\mathcal{G}_t)}{\mathbb{P}(\tau_A > t, \tau_B > t|\mathcal{G}_t)} \\ &= 1_{\{\tau_A > t, \tau_B > t\}} E(e^{-\int_t^s [\lambda_B(u) - \alpha\lambda_B(u)\int_0^t \lambda_A(v)dv]du}|\mathcal{G}_t)\end{aligned} \tag{2.12}$$

Comparing the above equation with equation (2.10), we see that there has a term $\alpha\lambda_B(u)\int_0^t \lambda_A(v)dv$ subtracted from bond B's intensity function when its counterparty bond A has not defaulted up to time t.

2.5 Our Dependent Default Risk Model

From the above section, we know that one firm's default will cause its counterparty's intensity function to be added a term. As in Jarrow & Yu [31], we also use indicator function to incorporate the event correlation. However, we don't assume that the default effect will be always on at its counterparty's rest time period. Instead, we assume that the holding time of this default effect follows exponential distribution ($\eta, \mathbb{P}(\eta > t) = e^{-\mu t}$, and its density function $f_\eta(t) = \frac{d\mathbb{P}_\eta(t)}{dt} = \mu e^{-\mu t}$). It means that the default effect will disappear after a certain time period. Now, we let $\lambda_A = a > 0$, so its density function $f_{\tau_A}(t) = ae^{-at}$, and let

$$\lambda_B(t) = b_1 + b_2 1_{\{\tau_A \leqslant t \leqslant \tau_A + \eta\}} \tag{2.13}$$

where η is exponential distribution with parameter μ, and we assume that the default time τ_A and η are independent. Now, we want to find the firm B's survival probability based on whether firm A has defaulted or not up to time t.

Case I: firm A has already defaulted by time t, $1_{\{\tau_A \leqslant t\}} = 1$.

Since firm A has already defaulted by time t, so we know the default time τ_A, say $\tau_A = S$. Here, S is deterministic and $0 \leqslant S \leqslant t$. For $u \geqslant t$, we have $(\tau_A \leqslant t) \subset (\tau_A \leqslant u)$ and $1_{\{\tau_A \leqslant t\}} = 1_{\{\tau_A \leqslant u\}} = 1$. So

$$\lambda_B(u) = b_1 + b_2 1_{\{\tau_A \leqslant u \leqslant \tau_A + \eta\}} = b_1 + b_2 1_{\{S \leqslant u \leqslant S + \eta\}} = b_1 + b_2 1_{\{u \leqslant S + \eta\}}$$

Therefore, the survival probability of firm B is:

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(u)du)|\mathcal{G}_t)$$

$$= E(\exp(-\int_t^T [b_1 + b_2 1_{\{u \leqslant S+\eta\}}]ds)|\mathcal{G}_t) \tag{2.14}$$

$$= e^{-b_1(T-t)}\left\{1 - \frac{b_2}{b_2+\mu}e^{-\mu(t-S)} + \frac{b_2}{b_2+\mu}e^{-b_2(T-t)-\mu(T-S)}\right\}$$

Check appendix to see the derivation of equation (2.14). In this case, the time-t price of zero-coupon bonds issued by B with maturity T with recovery of treasury assumption is:

$$v_B(t,T) = E(\delta_B 1_{\{\tau_B \leqslant T\}} e^{-\int_t^T r(s)ds} + 1_{\{\tau_B > T\}} e^{-\int_t^T r(s)ds}|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E(\delta_B e^{-\int_t^T r(s)ds} + (1-\delta_B)1_{\{\tau_B > T\}} e^{-\int_t^T r(s)ds}|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= \delta_B P(t,T) + (1-\delta_B)P(t,T)E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t) \tag{2.15}$$

$$= \delta_B P(t,T) + (1-\delta_B)P(t,T)e^{-b_1(T-t)}\left\{1 - \frac{b_2}{b_2+\mu}e^{-\mu(t-S)}\right.$$

$$\left. + \frac{b_2}{b_2+\mu}e^{-b_2(T-t)-\mu(T-S)}\right\}$$

We know that the expectation of η is $\frac{1}{\mu}$. It means that the holding time of the default effect will have $\frac{1}{\mu}$ much long. So the smaller value of μ, the longer holding time of default effect. In Jarrow & Yu [31], they let the default effect be alive all the rest of its counterparty life time. So when μ gets smaller and smaller, both

results from our model and Jarrow & Yu [31] should be closer and closer.

Note: (a) A quick check to find that when there is no event correlation (i.e. $b_2 = 0$), then equation (2.14) goes to $e^{-b_1(T-t)}$. i.e. ,

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = e^{-b_1(T-t)}$$

It means that the bond A 's default information has no effect on bond B's survival probability.

(b) Let $\mu \to 0$, then equation (2.14) converges to $e^{-(b_1+b_2)(T-t)}$. So we have

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t) = e^{-(b_1+b_2)(T-t)}$$

This is exactly the same result as in Jarrow & Yu [31].

(c) Let $\mu \to \infty$, then equation (14) converges to $e^{-b_1(T-t)}$. i.e.,

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = e^{-b_1(T-t)}$$

It means that instantaneous default effect also has no effect on bond B's survival probability.

Case II: firm A has not defaulted up to time t, $1_{\{\tau_A > t\}} = 1$.

Now, we know that $\tau_A > t$, and for $s > t$, $\mathbb{P}(\tau_A > s|\mathcal{G}_t) = e^{-a(s-t)}$. Using the property $E(X) = E(E(X|Y))$, where X, Y are random variables. Then the survival

probability of firm B is:

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t)$$

$$= e^{-b_1(T-t)} \int_0^\infty [E_t(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + y\}} ds)|\eta = y)]\mu e^{-\mu y} dy \tag{2.16}$$

where E_t denotes the expectation given the information filtration $\mathcal{G}_t$. Appendix gives the proves of the following three equations.

When $a = b_2, a \neq b_2 + \mu$, then

$$E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t)$$

$$= e^{-b_1(T-t)}\{\frac{\mu}{a+\mu} + \frac{a}{\mu}e^{-a(T-t)} - \frac{a^2}{\mu(\mu+a)}e^{-(a+\mu)(T-t)}\} \tag{2.17}$$

When $a \neq b_2, a \neq b_2 + \mu$, then

$$E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t)$$

$$= e^{-b_1(T-t)}\left\{\frac{\mu}{b_2+\mu} - \frac{b_2}{a-b_2-\mu}e^{-a(T-t)} + \frac{ab_2}{(b_2+\mu)(a-b_2-\mu)}e^{-(b_2+\mu)(T-t)}\right\}. \tag{2.18}$$

When $a \neq b_2, a = b_2 + \mu$, then

$$E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t)$$

$$= e^{-b_1(T-t)}\left\{\frac{\mu}{b_2+\mu} + \frac{b_2(1-\mu(T-t))}{b_2-a}e^{-a(T-t)} - \frac{b_2(\mu+a)}{(b_2-a)(b_2-\mu)}e^{-(b_2+\mu)(T-t)}\right\}. \tag{2.19}$$

In this case, the time-t price of zero-coupon bonds issued by B with maturity T with recovery of treasury assumption is:

$$v_B(t,T) = E(\delta_B 1_{\{\tau_B \leqslant T\}} e^{-\int_t^T r(s)ds} + 1_{\{\tau_B > T\}} e^{-\int_t^T r(s)ds} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E(\delta_B e^{-\int_t^T r(s)ds} + (1-\delta_B) 1_{\{\tau_B > T\}} e^{-\int_t^T r(s)ds} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \tag{2.20}$$

$$= \delta_B P(t,T) + (1-\delta_B) P(t,T) E(exp(-\int_t^T \lambda_B(s)ds) | \mathcal{G}_t)$$

Now using one of equations (2.17), (2.18) and (2.19) to substitute into the expectation in the above equation, we can find the time-t price of $v_B(t,T)$.

Theoretically, in this case, when μ is getting bigger and bigger, the value $\frac{v_B(t,T)}{p(t,T)}$ from our model is getting closer and closer to $e^{-b_1(T-t)}$. It means that when the default effect of A on B can't hold a certain period, then it'll has no effect on B though A and B are highly correlated.

Note: (a) It's easy to check that equation (2.17) goes to $(a(T-t)+1)e^{-(a+b_1)(T-t)}$ as μ goes to 0. And both equations (2.18) and (2.19) go to $\frac{b_2 e^{-(b_1+a)(T-t)} - a e^{-(b_1+b_2)(T-t)}}{b_2 - a}$ as μ goes to 0.

(b) When $b_2 = 0$ (i.e. no event correlation), then all three equations (2.17), (2.18), and (2.19) are equal to $e^{-b_1(T-t)}$. So equation (2.16) can be simplified as follow:

$$\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = e^{-b_1(T-t)}$$

c) As μ goes to ∞, all three equations (2.17), (2.18), and (2.19) converge to $e^{-b_1(T-t)}$. As in the previous case, the instantaneous default effect has no effect on bond B's survival probability.

Now we'll let the default probability be dependent on interest rate, i.e. we let the bond A and bond B's intensity function be function of interest rate.

$$\lambda_A(t) = a_0 + a_1 r(t)$$

And

$$\lambda_B(t) = b_0 + b_1 r(t) + b_2 1_{\{\tau_A \leqslant t \leqslant \tau_A + \eta\}}$$

Here we just give out the survival probability $\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$ of bond B. We assumed that η is independent with $r(t)$.

When bond A has defaulted by time t, S is the default time of bond A.

Using equation (2.14) and the independent assumption of η and $r(t)$, we have:

$$\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T [b_0 + b_1 r(u) + b_2 1_{\{\tau_A \leqslant u \leqslant \tau_A + \eta\}}] du) | \mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(\exp(-\int_t^T [b_1 r(u) + b_2 1_{\{u \leqslant S + \eta\}}] du) | \mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(\exp(-b_1 \int_t^T r(u) du) | \mathcal{G}_t) E(\exp(-b_2 \int_t^T 1_{\{u \leqslant S+\eta\}}) | \mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(e^{-b_1 R_{t,T}} | \mathcal{G}_t) \left\{ 1 - \frac{b_2}{b_2 + \mu} e^{-\mu(t-S)} + \frac{b_2}{b_2 + \mu} e^{-b_2(T-t) - \mu(T-S)} \right\} \tag{2.21}$$

Where $R_{t,T} = \int_t^T r(u) du$.

When bond A has not defaulted up to time t.

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t)$$

$$= e^{-b_0(T-t)}E(\{e^{-(b_2+\mu)(T-t)-b_1R_{t,T}}(1+b_2\int_t^T e^{-(a_0-b_2)(x-t)-a_1R_{t,x}}dx)$$

$$+\frac{\mu}{\mu+b_2}e^{-b_1R_{t,T}} + \mu b_2 e^{-b_2T+a_0t-b_1R_{t,T}}\int_0^{T-t}\int_{T-y}^T [e^{(b_2-a_0)x-a_1R_{t,x}}dx]e^{-\mu y}dy\}|\mathcal{G}_t). \tag{2.22}$$

2.5.1 Numerical Examples

For simplicity, we assume the recovery rate δ_B is zero. As in Yu [58], we use the normalized (against the default-free bond price) zero-coupon bond prices (i.e. $\frac{v_B(t,T)}{p(t,T)}$) to compare our model with Jarrow & Yu [31].

μ	Wang				Jarrow & Yu			
	$b_2 = 0.5$		$b_2 = 5$		$b_2 = 0.5$		$b_2 = 5$	
	T=2	T=11	T=2	T=11	T=2	T=11	T=2	T=11
5000	-9.9E-05	-9.5E-04	-9.9E-04	-9.5E-03	-0.212	-7.68	-0.798	-9.33
50	-9.7E-03	-0.094	-0.089	-0.864	-0.212	-7.68	-0.798	-9.33
5	-0.074	-0.85	-0.448	-4.7	-0.212	-7.68	-0.798	-9.33
1	-0.16	-2.97	-0.692	-7.8	-0.212	-7.68	-0.798	-9.33
0.5	-0.183	-4.30	-0.741	-8.5	-0.212	-7.68	-0.798	-9.33
0.1	-0.206	-6.66	-0.786	-9.16	-0.212	-7.68	-0.798	-9.33
0.01	-0.212	-7.68	-0.798	-9.33	-0.212	-7.68	-0.798	-9.33

Table 2.1. Percentage change of values from our model vs. values when there is no counterparty risk (when $t = 1, a = 0.01, b_1 = 0.01,$ $\boxed{b_2 = 0.5}$ and no default up to time t)

From the above table, we can see that the value which is from our model is very close to the value when there is no counterparty risk when μ is large even with the high correlation . For example, when $\mu = 5000, T - t = 1, b_2 = 5$, our value is $-(9.51E-03)\%$ less. It means that if the default effect is instantaneous, then this default event will have no effect on its counterparty. This is intuitive. However, Jarrow & Yu model can't reflect this phenomena.

μ	$b_2 = 0.5$		$b_2 = 5$	
	T-t=1	T-t=10	T-t=1	T-t=10
5000	0.21268	8.322	0.803	10.286
50	0.2031	8.221	0.715	9.343
5	0.138	7.402	0.353	5.908
1	0.0523	5.106	0.107	1.689
0.5	0.0291	3.664	0.057	0.919
0.01	6.512E-04	0.1256	1.233E-03	0.0201
0.001	6.526E-05	0.0127	1.224E-03	2.019E-03

Table 2.2: Percentage change of values from our model vs. Jarrow & Yu [31] (when $t = 1, a = 0.01, b_1 = 0.01$, and no default up to time t)

The above table tells us that when μ is small ,then results from the different models are very close no matter how much correlation ($b_2 = 0.5, 5$) the two bonds have and how long the date-to-maturity ($T - t = 1, 10$) is. For example, for $\mu = 0.001$, the difference of bond prices from the two models is only 0.0127% when $T - t = 10$. However, when $\mu = 50, b_2 = 0.5$ is large and the date-to-maturity is relative long ($T - t = 10$), then the bond price from our model will be 8.221% larger than that of Jarrow & Yu.

b_2	Wang	Jarrow & Yu	Percentage change
0.02	0.9231276	0.9231163	1.22E-03
0.1	0.6441293	0.6440364	0.0144
0.4	0.4318622	0.4317105	0.0351
0.6	0.2895725	0.2893842	0.0606
0.8	0.1941904	0.1939800	0.1085
1	0.130219	0.1300287	0.1716

Table 2.3: Percentage change of values from our model vs. values from Jarrow & Yu (when $t = 3, S = 1, b_1 = 0.02, T = 5$, $\boxed{\mu = 0.0001}$, and default has occurred by time t)

From the above table, we can see that when the holding time of default effect is long (i.e. μ is small, $\mu = 0.0001$), then results from our model is almost the same as that of Jarrow & Yu [2001] no matter how much correlation the two firms have. For example, when $b_2 = 1$, the bond prices from our model is only 0.1716% larger than that of Jarrow & Yu [31].

b_2	Wang	NCR	Percentage change
1	0.9607886	0.96078944	-8.777E-05
2	0.9607879	0.96078944	-1.536E-04
3	0.9607897	0.96078944	-2.448E-04
4	0.9607875	0.96078944	-2.458E-04
5	0.9607871	0.96078944	-2.793E-04
10	0.9607858	0.96078944	-3.84E-04
20	0.9607849	0.96078944	-4.726E-04

Table 2.4: Percentage change of values from our model vs. values from NCR (when $t = 3, S = 1, b_1 = 0.02, T = 5, \boxed{\mu = 6}$, and default has occurred by time t)

From the above table, we can see that when we choose a little bit larger value $\mu = 6$ (i.e. with relative short holding time of default effect, the default effect will last two months), then though the correlation becomes stronger (i.e. b_2 increases), the difference of the bond prices from the different models is still very small. For example, for $b_2 = 20$, our bond price is only 4.726E-04% smaller than the bond price from NCR.

Next table will show the percentage change of values from our model vs. the values of Jarrow & Yu, and percentage change of values from model vs. the values when there is no counterparty risk as the length of holding time of default effect (μ) changes. We pick some specific numbers for $b_1, b_2, 1, S, T$. We let $b_1 = 0.02, b_2 = 0.02, S = 1, t = 3, T = 11$. From Jarrow & Yu model, the normalized bond price is 0.726149037. The normalized bond price is 0.852143789 when there is no counterparty risk (NCR).

μ	Wang	Wang vs. Jarrow & Yu	Wang vs. NCR
0.01	0.0.73328	0.9887	-13.943
0.05	0.7576809	4.3423	-11.085
0.09	0.7764092	6.9215	-8.888
0.1	0.7803387	7.4693	-8.421
0.2	0.8091497	11.430	-5.045
0.3	0.8251742	13.637	-3.165
0.4	0.8345441	14.927	-2.065
0.5	0.8402748	15.717	-1.393
0.6	0.8439225	16.219	-0.965
0.7	0.8463251	16.55	-0.683
1	0.8498832	17.04	-0.265

Table 2.5: Percentage change of values from our model vs. values from Jarrow & Yu, and values with no counterparty risk (when default has occurred by time t)

From table 2.5, we can see that

(a) when $\mu > \frac{1}{t-S} = 0.5$, it means that the default effect disappeared before the current time $t = 3$. So the normalized bond prices from our model has no significant difference from that of NCR. For example, when $\mu = 1$, the normalized bond price from our model is only 0.265% smaller than that of NCR.

(b) When $\mu < \frac{1}{T-S} = 0.1$, it means that the default effect will last all the rest of bond B's life. So the normalized bond price from our model has no significant difference from that of Jarrow & Yu. For example, when $\mu = 0.01$, the normalized bond price from our model is only 0.9887% higher than that of Jarrow & Yu.

(c) When $0.1 = \frac{1}{T-S} < \mu < \frac{1}{t-S} = 0.5$, it means that the default effect is still alive at current time t, but it'll disappear before the bond B's maturity date T. In this

case, the normalized bond price from our model has significant difference form that of Jarrow & Yu. For example, when $\mu = 0.2$, the normalized bond price from our model will be 11.43% higher than that of Jarrow & Yu.

2.5.2 Conclusion

We introduce the length of holding time of default effect into Jarrow & Yu [31] model. We find that when we assume long period of holding time (small μ), then results from our model and Jarrow & Yu model are very close. However, when two firms are highly correlated and one firm's default effect will not last long, then results from the two models are quite different. Our model can reflect the intuitive phenomena that if the firm can survive, then the default effect from its counterparty will disappear as time goes.

2.6 Valuation of Default Swap of the First-to-Default Baskets

2.6.1 Introduction

First-to-default baskets are credit derivatives which are based on a portfolio of underlying reference entities. Rather than holding default risk on many individual credits, buyers of the first-to-default protection can pool the credit in a basket. The buyer of protection in this structure is hedged against the risk of default only with this basket of reference entities. A default swap is a type of default insurance. The buyer of the default protection makes a regular payment quoted as a percentage of the notional amount per year which is called swap rate premium. These payments continue until either the expiration of the swap or a default event by the underlying reference. If there is a default, the protection buyer delivers to the protection seller bonds of the reference entities and receives par value.

First-to-default baskets are used to hedge against credit risk for many reasons:

(a) the protection buyer gains on an entire portfolio of exposure.

(b) it is typically cheaper to buy a first-to-default basket than to purchase protection each credit individual credit.

(c) payout occurs when the first reference entity defaults.

The existing literature investigating the valuation of default swap of first-to-default baskets (see Li [42] and Kijima [35]) assumes that the default times of the underlying references $(\tau_1, \tau_2, \cdots, \tau_n)$ are independent, or conditional independent (also see Jarrow & Yu [31]) . However, we find that if the references' intensity functions of their default times are assumed to be functions of common factors (e.g. interest rate and market index), then the model will lose the event correlation as long as you use the conditional independent assumption, even you let the references' intensity functions be correlated (as in Kijima [35]). So we use the model derived in section 2.5. Under this model, we still can use the conditional independent assumption while not losing the event correlation.

2.6.2 Relationship between Event Correlation and Conditional Independent Assumption

Event correlation refers to how a firm's default probability is affected by default of other firms (see Jarrow & Turnbull [30]). Actually, when two firms are correlated, then one firm's default probability will be affected by its counterparty, even there is no default event (see Jarrow & Yu [31]). From Jeanblanc & Rutkowski [32], we know:

$$\mathbb{P}(\tau_i > s \big| \mathcal{G}_t \vee \mathcal{D}_t^j) = 1_{\{\tau_j > t\}} \frac{\mathbb{P}(\tau_i > s, \tau_j > t \big| \mathcal{G}_t)}{\mathbb{P}(\tau_j > t \big| \mathcal{G}_t)} + 1_{\{\tau_j \leqslant t\}} \frac{\mathbb{P}(\tau_i > s, \tau_j \leqslant t \big| \mathcal{G}_t)}{\mathbb{P}(\tau_j \leqslant t \big| \mathcal{G}_t)} \quad (2.23)$$

If firm i and firm j are independent, then $\mathbb{P}(\tau_i > t_i, \tau_j > t_j) = \mathbb{P}(\tau_i > t_i, \mathbb{P}(\tau_j > t_j)$. So

$$\mathbb{P}(\tau_i > t|\mathcal{G}_t \vee \mathcal{D}_t^j) = \mathbb{P}(\tau_i > t|\mathcal{G}_t)$$

It means that the information of the firm j's default or survival has no effect on the firm i's default probability. However, if firm i and firm j are correlated, then from equation we know that the information of firm j's default or survival up to time t will have effect on the firm i's default probability (see, section 2.5 for details). Now let's give out the definition of conditional independent assumption and discuss how this assumption will affect the event correlation. First of all, let's clarify the definition of default probability or survival probability first. According to Lando [40], the survival probability has the following equations:

$$\mathbb{P}(\tau_i > t|\mathcal{G}_t) = e^{-\int_0^t \lambda_i(u)du} \tag{2.24}$$

and

$$\mathbb{P}(\tau_i > s|\mathcal{G}_t) = E(E(1_{\{\tau_i > s\}}|\mathcal{G}_{T^*})|\mathcal{G}_t) = E(exp(-\int_0^s \lambda_i(u)du)|\mathcal{G}_t)$$

for $s \geqslant t$, and where $\mathcal{G}_t$ is the filtration generated by the information of state variables, but not including the information of default process. So if we include the information of default process $\mathcal{D}_t^i$,, then we have:

$$\mathbb{P}(\tau_i > t|\mathcal{G}_t \vee \mathcal{D}_t^i) = 1\{\tau_i > t\} \tag{2.25}$$

for $1\{\tau_i > t\}$ is $\mathcal{D}_t^i$-measurable. And if $s \geqslant t$, we have:

$$\mathbb{P}(\tau_i > s|\mathcal{G}_t \vee \mathcal{D}_t^i) = 1_{\{\tau_i > t\}} \frac{\mathbb{P}(\tau_i > s|\mathcal{G}_t)}{\mathbb{P}(\tau_i > t|\mathcal{G}_t)} = 1_{\{\tau_i > t\}} E(e^{-\int_t^s \lambda_i(u)du}|\mathcal{G}_t) \tag{2.26}$$

From equation (2.25), we know that the default has occurred or not at time t. Equation (2.26) tells us that we don't know when the default will occur in the future. We know that most intensity-based credit risk models (see, Lando [40], Jarrow & Turnbull [30], Jarrow & Yu [31], Jeanblanc & Rutkowski [32], etc.) have

the above features. Now we'll find that what realization of filtration the conditional independent assumption uses is crucial to keep the event correlation in the model.

Conditional independent assumption means that given the realization of filtration, the default times of the n underlying references are independent, i.e. we have:

$$\begin{aligned} &\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2, \cdots, \tau_n > t_n | \mathcal{J}_{T^*}) \\ &= \mathbb{P}(\tau_1 > t_1 | \mathcal{J}_{T^*}) \mathbb{P}(\tau_2 > t_2 | \mathcal{J}_{T^*}) \cdots \mathbb{P}(\tau_n > t_n | \mathcal{J}_{T^*}) \end{aligned} \tag{2.27}$$

If we choose $\mathcal{J}_{T^*} = \mathcal{G}_{T^*}$ (as in Jarrow & Yu [31], Kijima [35]), then the event correlation will be eliminated by the conditional independent assumption. We can explain this by both the marginal distribution of default time and their joint distribution. For simplicity, we just consider the case when $n = 2$, so conditional independent assumption has the form:

$$\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2 | \mathcal{G}_{T^*}) = \mathbb{P}(\tau_1 > t_1 | \mathcal{G}_{T^*}) \mathbb{P}(\tau_2 > t_2 | \mathcal{G}_{T^*})$$

where $t < t_1 \leqslant T^*, t < t_2 \leqslant T^*$.

The marginal distribution becomes:

$$\mathbb{P}(\tau_1 > t_1|\mathcal{G}_t \vee \mathcal{D}_t^2) = 1_{\{\tau_2>t\}}\frac{\mathbb{P}(\tau_1 > t_1, \tau_2 > t\big|\mathcal{G}_t)}{\mathbb{P}(\tau_2 > t\big|\mathcal{G}_t)} + 1_{\{\tau_2\leqslant t\}}\frac{\mathbb{P}(\tau_1 > t_1, \tau_2 \leqslant t\big|\mathcal{G}_t)}{\mathbb{P}(\tau_2 \leqslant t\big|\mathcal{G}_t)}$$

$$= 1_{\{\tau_2>t\}}\frac{E(\mathbb{P}(\tau_1 > t_1, \tau_2 > t\big|\mathcal{G}_{T^*})|\mathcal{G}_t)}{\mathbb{P}(\tau_2 > t\big|\mathcal{G}_t)} + 1_{\{\tau_2\leqslant t\}}\frac{E(\mathbb{P}(\tau_1 > t_1, \tau_2 \leqslant t\big|\mathcal{G}_{T^*})|\mathcal{G}_t)}{\mathbb{P}(\tau_2 \leqslant t\big|\mathcal{G}_t)}$$

$$= 1_{\{\tau_2>t\}}\frac{E(\mathbb{P}(\tau_1 > t_1|\mathcal{G}_{T^*})\mathbb{P}(\tau_2 > t\big|\mathcal{G}_{T^*})|\mathcal{G}_t)}{\mathbb{P}(\tau_2 > t\big|\mathcal{G}_t)}$$

$$+ 1_{\{\tau_2\leqslant t\}}\frac{E(\mathbb{P}(\tau_1 > t_1|\mathcal{G}_{T^*})\mathbb{P}(\tau_2 \leqslant t\big|\mathcal{G}_{T^*})|\mathcal{G}_t)}{\mathbb{P}(\tau_2 \leqslant t\big|\mathcal{G}_t)} \tag{2.28}$$

Since $\mathbb{P}(\tau_2 > t|\mathcal{G}_{T^*}) = \mathbb{P}(\tau_2 > t|\mathcal{G}_t) = e^{-\int_0^t \lambda_2(u)du}$ and $e^{-\int_0^t \lambda_2(u)du}$ is $\mathcal{G}_t$-measurable, so the above equation reduces to:

$$\mathbb{P}(\tau_1 > t_1|\mathcal{G}_t \vee \mathcal{D}_t^2) = \mathbb{P}(\tau_t > t_1|\mathcal{G}_t)$$

The above equation tells us that the information of default process $\mathcal{D}_t^2$ has no effect on its counterparty's survival probability. Now let's check its joint distribution $\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2)$. Use Farlie-Gumbel-Morgenstern copula $C(u, v) = uv(1 + \alpha(1 - u)(1 - v)), \alpha \in [-1, 1]$, we know that:

$$\begin{aligned}&\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2)\\ &= E(e^{-\int_0^{t_1} \lambda_1(u)du}e^{-\int_0^{t_2} \lambda_2(u)du}[1 + \alpha(1 - e^{-\int_0^{t_1} \lambda_1(u)du})(1 - e^{-\int_0^{t_2} \lambda_2(u)du})])\end{aligned} \tag{2.29}$$

Using the conditional independent assumption, we get:

$$\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2) = E(e^{-\int_0^{t_1} \lambda_1(u)du}e^{-\int_0^{t_2} \lambda_2(u)du}) \tag{2.30}$$

Comparing the above two equations, we know that the result of the conditional independent assumption is the same as we assume $\alpha = 0$ in the Farlie-Gumbel-Morgenstern copula. Therefore, there'll be no event correlation in the model.

2.6.3 Valuation of the Default Swap

Let $\tau = \min_{1 \leqslant i \leqslant n} \tau_i$ be the first-to-default time and T is the maturity date of the default swap. And all risky bonds' maturities are longer than the default swap's maturity, i.e. $T < \min_{1 \leqslant i \leqslant n}(T_i)$. We assume that both the protection seller and protection buyer are default-free. And the swap rate premium is paid at rate U from protection buyer to protection seller at time t_j, $j = 1, 2, \cdots, m$

$$t \leqslant t_1 < t_2 < \cdots < t_m = T$$

If default occurs during $(t_k, t_{k+1}]$, then the payment terminates at time t_k. We assume that the n risky bonds in the basket are alive at time t, i.e. $\tau_i > t$. So the time-t value of the payment, denoted by $V_{pb}1_{\tau>t}$, from protection buyer to protection seller is

$$V_{pb}1_{\{\tau>t\}} = \sum_{j=1}^{m} E(U1_{\{\tau>t_j\}}e^{-\int_t^{t_j} r(u)du}|\mathcal{F}_t)$$

On the other hand, if default occurs before the default swap maturity date T and the recovery rate of risky bond i is δ_i, under the assumption of recovery of treasury, then the time-t value of the payment from protection seller to protection buyer is

$$V_{ps}1_{\{\tau>t\}} = E(e^{-\int_t^{\tau} r(u)du} \sum_{i=1}^{n}(1-\delta_i)1_{\{\tau=\tau_i \leqslant T_{t_j}\}}|\mathcal{F}_t)$$

So the swap value at time t (denoted by V_t) to the protection seller is

$$\begin{aligned} V_t &= V_{pb}1_{\{\tau>t\}} - V_{ps}1_{\{\tau>t\}} \\ &= \sum_{j=1}^{m} E(U1_{\{\tau>t_j\}}e^{-\int_t^{t_j} r(u)du}|\mathcal{F}_t) - E(e^{-\int_t^{\tau} r(u)du} \sum_{i=1}^{n}(1-\delta_i)1_{\{\tau=\tau_i \leqslant T_{t_j}\}}|\mathcal{F}_t) \end{aligned} \tag{2.31}$$

If the default swap contract is initialized at time t, then the swap value at time t to the protection seller should be 0, i.e. $V_{pb} = V_{ps}$. So the default swap rate is determined by the following formula.

$$U = \frac{E(e^{-\int_t^\tau r(u)du} \sum_{i=1}^n (1-\delta_i) 1_{\{\tau=\tau_i \leqslant T_{t_j}\}} | \mathcal{F}_t)}{\sum_{j=1}^m E(1_{\{\tau > t_j\}} e^{-\int_t^{t_j} r(u)du} | \mathcal{F}_t)} \tag{2.32}$$

In the following part of this section, we want to calculate the swap rate premium U.

For simplicity, we just consider when $n = 2$, say bond A and bond B which are correlated. As in Jarrow & Yu [31], we treat bond A as a primary bond, bond B as a secondary bond. It also means that if bind A gets default, then the intensity of bond B will be added one term. However, if bond B gets default, there is no effect on bond A. We use τ_A, τ_B to denote their default times, and $\lambda_A(t), \lambda_B(t)$ to denote their intensity processes, and $\mathcal{D}_t^A = \sigma(1_{\{\tau_A \leqslant s\}}, s \leqslant t), \mathcal{D}_t^B = \sigma(1_{\{\tau_B \leqslant s\}}, s \leqslant t)$. Now we'll analyze how the event correlation and holding time of default effect on the default swap premium. Our philosophy is to introduce the simplest model that will capture the event correlation and holding time of default effect. We use the model developed in section 2.5 to deal with the event correlation. Meanwhile, we assume that the interest rate $r(t)$ is a constant ($r(t) = r$, the recovery rate $\delta_A = \delta_B = 0$, bond A and bond B have same maturity date T_0. For completeness, we give out model from section 2.5 and results here. Let $\lambda_A(t) = a > 0$, and

$$\lambda_B(t) = b_1 + b_2 1_{\{\tau_A \leqslant t \leqslant \tau_A + \eta\}} \tag{2.33}$$

where η controls the holding time of bond A's default effect on bond B with the law of exponential distribution with parameter μ.

Case I: firm A has already defaulted by time t, $1_{\{\tau_A \leqslant t\}} = 1$.

The survival probability of firm B is:

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(exp(-\int_t^T \lambda_B(u)du)|\mathcal{G}_t)$$

$$= E(exp(-\int_t^T [b_1 + b_2 1_{\{u \leqslant S+\eta\}}]ds)|\mathcal{G}_t) \tag{2.34}$$

$$= e^{-b_1(T-t)}\left\{1 - \frac{b_2}{b_2+\mu}e^{-\mu(t-S)} + \frac{b_2}{b_2+\mu}e^{-b_2(T-t)-\mu(T-S)}\right\}$$

Where S is bond A's the default time.

Case II: firm A has not defaulted up to time t, $1_{\{\tau_A > t\}} = 1$.

When $a \neq b_2, a \neq b_2 + \mu$, then

$$E(exp(-\int_t^s \lambda_B(u)du)|\mathcal{G}_t)$$

$$= e^{-b_1(s-t)}\left\{\frac{\mu}{b_2+\mu} - \frac{b_2}{a-b_2-\mu}e^{-a(s-t)} + \frac{ab_2}{(b_2+\mu)(a-b_2-\mu)}e^{-(b_2+\mu)(s-t)}\right\}. \tag{2.35}$$

When $a = b_2, a \neq b_2 + \mu$, then

$$E(exp(-\int_t^s \lambda_B(u)du)|\mathcal{G}_t) \tag{2.36}$$

$$= e^{-b_1(s-t)}\{\frac{\mu}{a+\mu} + \frac{a}{\mu}e^{-a(s-t)} - \frac{a^2}{\mu(\mu+a)}e^{-(a+\mu)(s-t)}\}$$

When $a \neq b_2, a = b_2 + \mu$, then

$$E(exp(-\int_t^s \lambda_B(u)du)|\mathcal{G}_t)$$

$$= e^{-b_1(s-t)}\left\{\frac{\mu}{b_2+\mu} + \frac{b_2(1-\mu(s-t))}{b_2-a}e^{-a(s-t)} - \frac{b_2(\mu+a)}{(b_2-a)(b_2-\mu)}e^{-(b_2+\mu)(s-t)}\right\}. \tag{2.37}$$

Now, let's simplify the denominator in equation (2.32). Using the conditional independent assumption, we have:

$$\mathbb{P}(\tau_A > t_j, \tau_B > t_j|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= \mathbb{P}(\tau_A > t_j|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)\mathbb{P}(\tau_B > t_j|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \tag{2.38}$$

$$= e^{-a(t_j-t)}e^{-b_1(t_j-t)} = e^{-(a+b_1)(t_j-t)}$$

And using the above equation, we have

$$E(e^{-\int_t^{t_j} r(u)du}1_{\{\tau>t_j\}}|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E([E(e^{-\int_t^{t_j} r(u)du}1_{\{\tau>t_j\}}|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)]|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \tag{2.39}$$

$$= E(e^{-(r+a+b_1)(t_j-t)}|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = e^{-(r+a+b_1)(t_j-t)}$$

Therefore the denominator in equation (2.32) is following:

$$\sum_{j=1}^m E(1_{\{\tau>t_j\}}e^{-\int_t^{t_j} r(u)du}|\mathcal{F}_t) = \sum_{j=1}^m e^{-(r+a+b_1)(t_j-t)} \tag{2.40}$$

For the numerator in equation (2.32), we have two cases: $\tau = \tau_B$ and $\tau = \tau_A$.

For the $\tau = \tau_B$ case, it'll be a little easier for we assume that the default of bond B has no effect on bond A. And we know that

$$\mathbb{P}(s - ds < \tau_B \leqslant s | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = b_1 e^{-b_1(s-t)} ds$$

And

$$\mathbb{P}(\tau_A > s, s - ds < \tau_B \leqslant s | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= \mathbb{P}(\tau_A > s | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \mathbb{P}(s - ds < \tau_B \leqslant s | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \tag{2.41}$$

$$= e^{-a(s-t)} b_1 e^{-b_1(s-t)} ds = b_1 e^{-(a+b_1)(s-t)} ds$$

Therefore, one term of numerator in equation (2.32) can be simplified as follow:

$$E(e^{-\int_t^{\tau_B} r(u)du} 1_{\{\tau = \tau_B \leqslant T_0\}} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E([E(e^{-\int_t^{\tau_B} r(u)du} 1_{\{\tau = \tau_B \leqslant T_0\}} | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)] | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E(\int_t^{T_0} e^{-r(s-t)} b_1 e^{-(a+b_1)(s-t)} ds | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \tag{2.42}$$

$$= \frac{b_1}{r + a + b_1} [1 - e^{-(r+a+b_1)(T_0 - t)}]$$

For the case when $\tau = \tau_A$, it'll be a little complicated for the default of bond A will affect the default probability of bond B. We know that

$$\mathbb{P}(s - ds < \tau_A \leqslant s | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = a e^{-a(s-t)} ds$$

And

$$\mathbb{P}(\tau_B > s | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = exp(-\int_t^s \lambda_B(u) du)$$

$$= e^{-b_1(s-t)} exp(-b_2 \int_t^s 1_{\{\tau_A \leqslant u \leqslant \tau_A + \eta\}} du) \tag{2.43}$$

Thus we have following for the other term of the numerator in equation (2.32).

$$E(e^{-\int_t^{\tau_A} r(u)du} 1_{\{\tau=\tau_A \leqslant T_0\}} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E([E(e^{-\int_t^{\tau_A} r(u)du} 1_{\{\tau=\tau_A \leqslant T_0\}} | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)] | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= E(\int_t^{T_0} e^{-r(s-t)} a e^{-(a+b_1)(s-t)} E(exp(-b_2 \int_t^s 1_{\{\tau_A \leqslant u \leqslant \tau_A+\eta\}} du)) ds | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \tag{2.44}$$

Now based on your specified values of a, μ, b_2, we can use equations (2.35),(2.37), (2.36) to substitute $E(exp(-b_2 \int_t^s 1_{\{\tau_A \leqslant u \leqslant \tau_A+\eta\}} du))$. After taking the integration, we have:

When $a \neq b_2, a \neq b_2 + \mu$, then

$$E(e^{-\int_t^{\tau_A} r(u)du} 1_{\{\tau=\tau_A \leqslant T_0\}} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$\begin{aligned} = \Bigg\{ & \frac{a\mu}{(b_2+\mu)(a+b_1+r)} (1 - e^{-(a+b_1+r)(T_0-t)}) \\ & + \frac{ab_2}{(a-b_2-\mu)(r+2a+b_1+\mu)} (e^{-(r+2a+b_1)(T_0-t)} - 1) \\ & + \frac{a^2 b_2}{(b_2+\mu)(a-b_2-\mu)(r+a+b_1+b_2+\mu)} (1 - e^{-(r+a+b_1+b_2+\mu)(T_0-t)}) \Bigg\} \end{aligned} \tag{2.45}$$

When $a \neq b_2, a = b_2 + \mu$, then

$$
\begin{aligned}
& E(e^{-\int_t^{\tau_A} r(u)du} 1_{\{\tau=\tau_A \leqslant T_0\}} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
& = \Bigg\{ \frac{ab_2}{(b_2-a)(2a+b_1+r)} (1 - e^{-(2a+b_1+r)(T_0-t)}) \\
& \quad + \frac{ab_2\mu}{(b_2-a)(2a+b_1+r)} \{ (T_0-t) e^{-(2a+b_1+r)(T_0-t)} \\
& \quad + \frac{1}{2a+b_1+r} (e^{-(2a+b_1+r)(T_0-t)} - 1) \} \\
& \quad + \frac{a\mu}{(b_2+\mu)(a+b_1+r)} (1 - e^{-(a+b_1+r)(T_0-t)}) \\
& \quad + \frac{ab_2(\mu+a)}{(b_2-a)(b_2-\mu)(a+r+b_1+b_2+\mu)} (e^{-(a+r+b_1+b_2+\mu)(T_0-t)} - 1) \Bigg\}
\end{aligned}
\tag{2.46}
$$

When $a = b_2, a \neq b_2 + \mu$, then

$$
\begin{aligned}
& E(e^{-\int_t^{\tau_A} r(u)du} 1_{\{\tau=\tau_A \leqslant T_0\}} | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
& = \Bigg\{ \frac{a\mu}{(a+\mu)(a+b_1+r)} (1 - e^{-(a+b_1+r)(T_0-t)}) \\
& \quad + \frac{a^2}{\mu(2a+b_1+r)} (1 - e^{-(2a+b_1+r)(T_0-t)}) \\
& \quad + \frac{a^3}{\mu(a+\mu)(2a+b_1+r+\mu)} (e^{-(2a+b_1+r+\mu)(T_0-t)} - 1) \Bigg\}
\end{aligned}
\tag{2.47}
$$

2.6.4 Numerical Examples

From the previous section, we can use equations (2.45),(2.46), (2.47) to find the default swap premium U. Now the basket consists of two defaultable zero-coupon bonds (Bond A and Bond B). we assume that both A and B have the same maturity date $T_0 = 10$ and their recovery rates are 0, the maturity of the default swap T is 2, and the default swap premium is paid semiannaully ($t_j = 0.5, 1.0, 1.5, 2.0$). We assume the current time $t = 0$, interest rate $r = 0.08$, $b_1 = 0.01$. Then we use different values of a, b_2, μ to describe different default probability, different correlation, and different length of holding time of default effect. We use Excel to get the following two tables.

μ	$b_0 = 0$	$b_1 = 0.1$		$b_2 = 1$		$b_2 = 10$	
	U	U	Change	U	Change	U	Change
0.001	0.03576	0.03558	-0.503	0.03518	-1.622	0.03505	-1.985
0.1	0.03576	0.03561	-0.419	0.03522	-1.51	0.03506	-1.957
1	0.03576	0.03570	-0.168	0.03544	-0.895	0.03511	-1.818
10	0.03576	0.03575	-0.028	0.03569	-0.196	0.03540	-1.007
100	0.03576	0.03576	0.0	0.03575	-0.028	0.03569	-0.196

Table 2.6: Default Swap Premiums for Different Correlations, when $a = 0.01$.

From the above table, we can see that the percentage change of default swap premiums between the case with correlation and the case with no event correlation is very small. For example, for $b_2 = 10$, its default swap premium is only 1.985% smaller than the value with no event correlation. So, when bond A's default probability is small (i.e. small a), then the event correlation has little effect on the default swap premium.

μ	$b_0 = 0$	$b_1 = 0.1$		$b_2 = 1$		$b_2 = 10$	
	U	U	Change	U	Change	U	Change
0.001	11.5591	11.4490	-1.23	10.6206	-8.37	7.7710	-32.957
0.1	11.5591	11.4510	-1.21	10.6356	-8.24	7.7957	-32.744
1	11.5591	11.4668	-1.07	10.7528	-7.23	8.0062	-30.924
10	11.5591	11.5215	-0.60	11.2038	-3.34	9.2807	-19.932
100	11.5591	11.5537	-0.32	11.5052	-0.74	11.0624	-4.56

Table 2.7: Default Swap Premiums for Different Correlations, when $a = 5$.

From the above table, we can see that when default probability of bond A is high (i.e. large a), then the event correlation does matter. For example, for fixed $\mu = 0.1$, as correlation becomes stronger (b_2 from 0.1 to 10), the percentage changes of default swap premiums gets bigger (from down 1.21% to down 32.957%). Meanwhile, we can also find that the length of holding time of default effect ($\frac{1}{\mu}$) is a significant factor on the default swap premium when the correlation is high. The longer the default effect holds, the smaller the default swap premium is. For example, for $a = 5, b_2 = 10$, if the length of holding effect is short ($\mu = 100$), then the default swap premium is only 4.56% smaller than the case with no correlation. However, if the length of holding time of default effect is long ($\mu = 0.001$), then the default swap premium will be 32.957% smaller than the case with no correlation.

2.6.5 Conclusion

Kijima [35] found that the correlations between bonds in the baskets have little effect on the default swap premium. However, we find that this case happens only when the bonds' default probabilities are very small. If there exists bond which has

high default probability in the basket, then the correlation is a key factor in pricing the default swap premium. Meanwhile, we also find that the length of holding time of default effect is also a key factor in pricing the default swap premium.

2.7 Valuation of Dependent Default Risk in Collateralized Bond Obligations

Recently, many new products are associated with credit risk portfolios. Examples of these include Collateralized Debt Obligations (CDOs) which include Collateralized Loan Obligations (CLOs) and Collateralized Bond Obligations (CBOs). The central idea of these trades is ratings arbitrage. Banks typically hold large groups of loans on their balance sheets. These loans will vary in degree of credit risk and the overall pool may or may not be well diversified. In a CDO the bank slices up these loans into various tranches which are rated by a rating agency and then sold on to investors. CBO is a multitranche debt structure which is similar to some respect to a Collateralized Mortgage Obligation (CMO) structure. Typically low-rated bonds rather than mortgage serve as the collateral. Interests and principal repayments received on the bond portfolio are passed through to owners of the derivative securities. However, these payments is contingent upon the time and identity of the first or second-to-default. Default dependency is indispensable to price/hedge these portfolios.

2.7.1 Introduction

Collateralized Debt obligations (CDOs) are a form of structurated finance used to securitize corporate bonds (collateralized bond obligations or CBOs) and bank loans (collateralized loan obligations or CLOs). With a CDO, assets are pooled in a portfolio and then rated securities are issued to fund the purchase of the assets. CDOs

can consist of a number of assets: loans, bonds, combination of loans and bonds, and a variety of other assets. In a collateralized debt obligation, a portfolio is created that contains approximately three tranches. All of the tranches reference the same securities. They are differentiated by levels of risk; the CDO issuer finds a buyer for each tranche depending on the buyer's particular risk appetite. As losses occur, they are incurred by the holders of the *First Loss Tranche, Mezzanine Tranche,* and *Senior Tranche* sequentially. The issuer pays the investor a fixed payment for the credit protection, and in return, the investor would make a payment to the issuer if one of the securities were to default when the cumulative default level falls within their tranche.

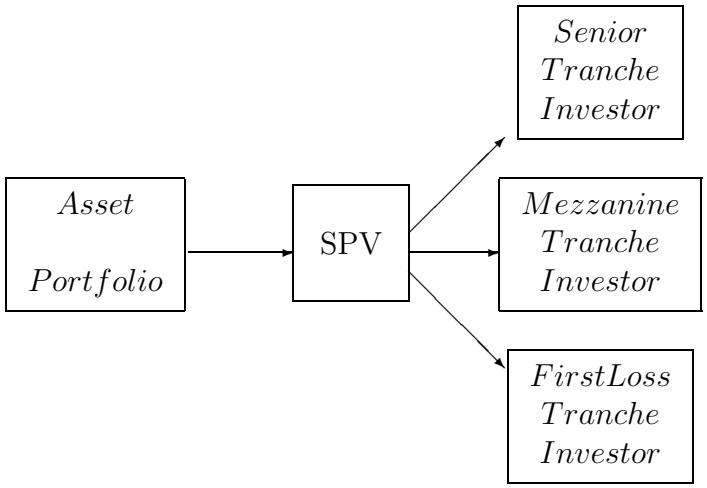

Figure 2.1: Typical CDO Structure

The above picture shows the typical CDO structure. The box labelled SPV denotes a 'Special Purpose Vehicle'. The SPV created for the issuance of a collateralized bond obligation will be a stand-lone, bankruptcy remote entity. For CBOs, it means that the asset portfolio is backed by high-yield corporate bonds. All subsequent payments made to CBO investors are derived from income received from the bond portfolio. The senior tranche investor and the mezzanine tranche investor receive

a specified coupon on their investment and are repaid their principal at maturity. These two tranches are also called bond tranches. They are first in priority of payment. The first loss tranche investor are paid receipts, with no guaranteed coupon or guarantee of principal repayment. That's why some papers call this tranche equity tranche. If the SPV actually receives all coupon and principal payments due from the bond portfolio, then all tranches investors will receive everything due to them and the first loss tranche investors will a receive a high return. If any of the collateral bonds defaults, then the first loss tranche investors will suffer a loss of return. This tranche investors are in effect making a leveraged investment in the high-yield portfolio. It is because of the redistribution of risk that the two bond tranches can obtain investment-grade ratings even though the underlying collateral consists largely of below investment-grade bonds.

A CBO is a correlation product. Investors in this product are buying correlation risk. To determine that they are getting a fair return for this risk, they must be able to measure the correlation risk. In this paper, we focus on two things:

(a) One is how the event correlation affects the credit protection of each tranche.

(b) The other is to analyze how the holding time of default effect affects the credit protection of each tranche.

2.7.2 Credit Risk Model

We just consider the simple case when there are only two bonds in the collateral pool, say bond A and bond B. As in Jarrow & Yu [31], we treat bond A as a primary bond, bond B as a secondary bond. It means that if bond B gets default,

there is no effect on bond A. However, if bond A gets default, then the intensity of bond B will be added one term. Using the model we just developed, we'll use exponential distribution to control the life of the added term. It means that after a certain time of period, this added term will disappear. We use τ_A, τ_B to denote their default times, and $\lambda_A(t), \lambda_B(t)$ to denote their intensity processes, and $\mathcal{D}_t^A = \sigma(1_{\{\tau_A \leqslant s\}}, s \leqslant t), \mathcal{D}_t^B = \sigma(1_{\{\tau_B \leqslant s\}}, s \leqslant t)$. We assume that the interest rate $r(t)$ is a constant ($r(t) = r$, bond A and bond B have same maturity date T_0. Let $\lambda_A(t) = a > 0$, and

$$\lambda_B(t) = b_1 + b_2 1_{\{\tau_A \leqslant t \leqslant \tau_A + \eta\}} \tag{2.48}$$

where $b_1 > 0, b_2 \geqslant 0$, and η controls the holding time of bond A's default effect on bond B with the law of exponential distribution with parameter μ. This extra term $b_2 1_{\{\tau_A \leqslant t \leqslant \tau_A + \eta\}}$ is induced by the default event of bond A. It will not appear before the bond A's default time τ_A and after the time $\tau_a + \eta$.

Case I: firm A has already defaulted by time t, $1_{\{\tau_A \leqslant t\}} = 1$.

The survival probability of firm B is:

$$\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(u)du) | \mathcal{G}_t)$$

$$= E(\exp(-\int_t^T [b_1 + b_2 1_{\{u \leqslant S+\eta\}}]ds) | \mathcal{G}_t) \tag{2.49}$$

$$= e^{-b_1(T-t)} \left\{ 1 - \frac{b_2}{b_2 + \mu} e^{-\mu(t-S)} + \frac{b_2}{b_2 + \mu} e^{-b_2(T-t)-\mu(T-S)} \right\}$$

Where S is bond A's default time ($S \leqslant t$).

Case II: firm A has not defaulted up to time t, $1_{\{\tau_A > t\}} = 1$.

When $a \neq b_2, a \neq b_2 + \mu$, then

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(u)du)|\mathcal{G}_t)$$

$$= e^{-b_1(T-t)}\left\{\frac{\mu}{b_2+\mu} - \frac{b_2}{a-b_2-\mu}e^{-a(T-t)} + \frac{ab_2}{(b_2+\mu)(a-b_2-\mu)}e^{-(b_2+\mu)(T-t)}\right\}. \tag{2.50}$$

When $a = b_2, a \neq b_2 + \mu$, then

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(u)du)|\mathcal{G}_t) \tag{2.51}$$

$$= e^{-b_1(T-t)}\{\frac{\mu}{a+\mu} + \frac{a}{\mu}e^{-a(T-t)} - \frac{a^2}{\mu(\mu+a)}e^{-(a+\mu)(T-t)}\}$$

When $a \neq b_2, a = b_2 + \mu$, then

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(u)du)|\mathcal{G}_t)$$

$$= e^{-b_1(T-t)}\left\{\frac{\mu}{b_2+\mu} + \frac{b_2(1-\mu(T-t))}{b_2-a}e^{-a(T-t)} - \frac{b_2(\mu+a)}{(b_2-a)(b_2-\mu)}e^{-(b_2+\mu)(T-t)}\right\}. \tag{2.52}$$

2.7.3 Credit Protection Valuation

Moody's uses a probabilistic, expected loss approach to determine a portfolio's credit risk. A portfolio's credit risk is quantified as the amount of loss protection needed to lower a portfolio's expected loss to the expected loss benchmark of the desired rating of the structured bonds, where expected loss is defined as the average of all possible principal losses weighted by their probability. For example, a single

speculative-grade bond with a 30% default probability and with a 70% loss of par, then its expected loss is $30\% \times 70\% = 21\%$. Similarly, the expected loss of the investment-grade bond with 5% default probability and a 70% loss of par is 3.5%. If the collateral pool is only backed by this speculative-grade bond, then the credit protection necessary for the portfolio to achieve an investment-grade rating is 58.3% which is calculated by the following formula:

$$\begin{aligned} &\text{Probability of Default } \times (\text{Default Severity} - \text{Credit Protection}) \\ &= \text{Target Expected Loss.} \end{aligned} \tag{2.53}$$

i.e. $30\% \times (70\% - X) = 3.5\% \Rightarrow X = 58.3\%$.

If the collateral pool has two bonds, then we can use the following formula to calculate credit protection:

$$\begin{aligned} &\text{Probability of 1 Bond Default } \times \\ &\text{(Default Severity if 1 Bond Defaults - Credit Protection)} \\ &+ \text{Probability of 2 Bond Default } \times \\ &\text{(Default Severity if 2 Bonds Default - Credit Protection)} \\ &= \text{Target Expected Loss.} \end{aligned} \tag{2.54}$$

If we assume that these two bonds are independent with same default probabilities (30%), and same default severity (70% loss of par), then plug in these numbers into the above formula, we get:

$$(2 \times 30\% \times 70\%) \times (35\% - X) + (30\% \times 30\%) \times (70\% - X) = 3.5\%.$$

Solving the above equation, we get the credit protection $X = 34.3\%$[3]. Comparing this number to that of only one speculative-grade bond in the collateral pool

[3]These two examples are from Lucas & McDaniel [1993].

(58.3%), we find that this one is much smaller. The reduction in credit protection is a result of the decreased variance in the portfolio's expected loss.

Now we use our dynamic credit risk model to calculate credit protection of the collateral pool. As we pointed previously, there are two correlated bonds A and B in the collateral pool. We assume the current time is t, and the maturity date of the CBO is T. So the probability of one bond default is $\mathbb{P}(\tau_A \leqslant T, \tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) + \mathbb{P}(\tau_A > T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$. And the probability of both bonds default is $\mathbb{P}(\tau_A \leqslant T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$. Let T^* denote the horizontal time of the economy (so $T \leqslant T^*$) and $\mathcal{G}_{T^*}$ denote the information filtration generated by the state variables (for example, interest rate, market index). Since bond B is a secondary bond, its default has no effect on bond A. It means that we have:

$$\mathbb{P}(\tau_A > T|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = e^{-\int_0^T \lambda_A(u)du} = e^{-a(T-t)}.$$

And

$$1_{\{\tau_A > T\}}\mathbb{P}(\tau_B \leqslant T|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = 1 - e^{-\int_0^T \lambda_B(u)du} = 1 - e^{-b_1(T-t)}.$$

Using conditional independent assumption[4], we have:

$$\begin{aligned}
&\mathbb{P}(\tau_A > T, \tau_B \leqslant T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= E(\mathbb{P}(\tau_A > T, \tau_B \leqslant T | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= E((\mathbb{P}(\tau_A > T | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)\mathbb{P}(\tau_B \leqslant T | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)) | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= e^{-a(T-t)}(1 - e^{-b_1(T-t)}).
\end{aligned} \tag{2.55}$$

To calculate the probability $\mathbb{P}(\tau_A \leqslant T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$,we know that $\mathbb{P}(\tau_A \leqslant T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = \mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) - \mathbb{P}(\tau_A > T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$. Using equations (2.50),(2.51), or (2.52), we can calculate $\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$. For the probability $\mathbb{P}(\tau_A > T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$, since bond A will survive up to time T, so there is no default

[4]Conditional independent assumption refers to the independence of default times τ_A and τ_B given the realization of information filtration. Here, we have

$$\mathbb{P}(\tau_A \in \mathcal{B}_1, \tau_B \in \mathcal{B}_2 | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = \mathbb{P}(\tau_A \in \mathcal{B}_1 | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)\mathbb{P}(\tau_B \in \mathcal{B}_2 | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B).$$

where $\mathcal{B}_1, \mathcal{B}_2$ are Borel sets.

effect. Therefore, we have:

$$\begin{aligned}
&\mathbb{P}(\tau_A > T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= E(\mathbb{P}(\tau_A > T, \tau_B > T | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= E(\mathbb{P}(\tau_A > T | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)\mathbb{P}(\tau_B > T | \mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= e^{-a(T-t)} e^{-b_1(T-t)}
\end{aligned} \tag{2.56}$$

So the probability of one bond default is following:

When $a \neq b_2, a \neq b_2 + \mu$, then

$$\begin{aligned}
&\mathbb{P}(\tau_A \leqslant T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) + \mathbb{P}(\tau_A > T, \tau_B \leqslant T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= e^{-a(T-t)} - 2e^{-(a+b_1)(T-t)} + e^{-b_1(T-t)} \Big\{ \frac{\mu}{b_2 + \mu} - \frac{b_2}{a - b_2 - \mu} e^{-a(T-t)} \\
&+ \frac{ab_2}{(b_2 + \mu)(a - b_2 - \mu)} e^{-(b_2+\mu)(T-t)} \Big\}.
\end{aligned} \tag{2.57}$$

When $a = b_2, a \neq b_2 + \mu$, then

$$\begin{aligned}
&\mathbb{P}(\tau_A \leqslant T, \tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) + \mathbb{P}(\tau_A > T, \tau_B \leqslant T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) \\
&= e^{-a(T-t)} - 2e^{-(a+b_1)(T-t)} + e^{-b_1(T-t)} \{ \frac{\mu}{a + \mu} + \frac{a}{\mu} e^{-a(T-t)} \\
&- \frac{a^2}{\mu(\mu + a)} e^{-(a+\mu)(T-t)} \}
\end{aligned} \tag{2.58}$$

When $a \neq b_2, a = b_2 + \mu$, then

$$\mathbb{P}(\tau_A \leqslant T, \tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) + \mathbb{P}(\tau_A > T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$$

$$= e^{-a(T-t)} - 2e^{-(a+b_1)(T-t)} + e^{-b_1(T-t)}\Bigg\{\frac{\mu}{b_2+\mu} + \frac{b_2(1-\mu(T-t))}{b_2-a}e^{-a(T-t)} \tag{2.59}$$

$$- \frac{b_2(\mu+a)}{(b_2-a)(b_2-\mu)}e^{-(b_2+\mu)(T-t)}\Bigg\}.$$

Using $\mathbb{P}(\tau_A \leqslant T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = \mathbb{P}(\tau_A \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) - \mathbb{P}(\tau_A \leqslant T, \tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$, and equation (2.50),(2.51), or (2.52), we can calculate the probability of two bonds default $\mathbb{P}(\tau_A \leqslant T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$. So we have:

When $a \neq b_2, a \neq b_2 + \mu$, then

$$\mathbb{P}(\tau_A \leqslant T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = 1 - e^{-a(T-t)} + e^{-(a+b_1)(T-t)}$$

$$- e^{-b_1(T-t)}\Bigg\{\frac{\mu}{b_2+\mu} - \frac{b_2}{a-b_2-\mu}e^{-a(T-t)} + \frac{ab_2}{(b_2+\mu)(a-b_2-\mu)}e^{-(b_2+\mu)(T-t)}\Bigg\}. \tag{2.60}$$

When $a = b_2, a \neq b_2 + \mu$, then

$$\mathbb{P}(\tau_A \leqslant T, \tau_B \leqslant T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = 1 - e^{-a(T-t)} + e^{-(a+b_1)(T-t)} \tag{2.61}$$

$$- e^{-b_1(T-t)}\{\frac{\mu}{a+\mu} + \frac{a}{\mu}e^{-a(T-t)} - \frac{a^2}{\mu(\mu+a)}e^{-(a+\mu)(T-t)}\}.$$

When $a \neq b_2, a = b_2 + \mu$, then

$$\mathbb{P}(\tau_A \leqslant T, \tau_B \leqslant T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = 1 - e^{-a(T-t)} + e^{-(a+b_1)(T-t)}$$

$$- e^{-b_1(T-t)} \left\{ \frac{\mu}{b_2 + \mu} + \frac{b_2(1 - \mu(T-t))}{b_2 - a} e^{-a(T-t)} - \frac{b_2(\mu + a)}{(b_2 - a)(b_2 - \mu)} e^{-(b_2+\mu)(T-t)} \right\}. \tag{2.62}$$

2.7.4 Numerical Examples

In this section, we'll analyze numerically how the event correlation (b_2) and the average of holding time of default effect ($\frac{1}{\mu}$) affect the credit protection. We assume the current time $t = 0$ and both bond A and bond B will lose 70% of par if default occurs. Now we use equations (2.57),(2.58), or (2.59), to calculate the probability of one bond default and use equations (2.60),(2.61), or (2.62) to calculate the probability of both bonds default. Here we assume CBO has the same maturity date as the bonds in the collateral pool. Then we choose $a = 0.0713, b_1 = 0.0713, T = 5$, thus both bond A and bond B has 30% default probability. Therefore we can compare our results to the one when bond A and bond B are independent (in this case, the credit protection is 34.3%).

b_2	Wang	Percentage change
0.01	0.3458678	0.81124
0.1	0.3676301	7.154
0.2	0.3862371	12.578
0.3	0.4005695	16.755
1	0.4462486	30.07
2	0.4643707	35.352

Table 2.8: Percentage changes of credit protections from our model vs. that of independent case. And $\boxed{\mu = 0.19}$.

In the above table we assume $\mu = 0.19$, the holding time of the default effect will last $\frac{1}{\mu} = 5.26$ years which is longer than the bonds' maturity date. The above table tells us that the credit protection increases as the correlation between the two bonds becomes strong. So the rating of the credit risk portfolio (i.e. collateral pool) is down. For example, when $b_2 = 2$, then the credit protection is 0.4643707 which is 35.352% larger than the one (0.34308) when both bonds are independent.

b_2	Wang	Percentage change
0.01	0.343088	1.15E-03
0.1	0.343124	0.0115
0.2	0.343163	0.023
0.3	0.343203	0.0345
1	0.343478	0.1147
2	0.343870	0.2288

Table 2.9: Percentage changes of credit protections from our model vs. that of independent case. And $\boxed{\mu = 365}$.

In table 2.9, we assume $\mu = 365$, so the holding time of the default effect is only 1 day. In this case, the correlation was not a key role in measuring the credit protection. For example, when $b_2 = 2$, the credit protection from our model is only 0.2288% higher than the one (0.34308) when both bonds are independent. In table 1, however, we get a credit protection which is 35.352% higher.

b_2	$\mu = 0.19$	$\mu = 0.25$	$\mu = 0.333$	$\mu = 0.5$	$\mu = 1$
0.01	0.811	0.745	0.667	0.546	0.343
0.1	7.154	6.613	5.969	4.954	3.205
0.2	12.578	11.70	10.645	8.962	5.973
0.3	16.755	15.674	14.366	12.249	8.384
1	30.07	28.875	27.365	24.747	19.193
2	35.352	34.49	33.364	31.305	26.415

Table 2.10: Percentage changes of credit protections from our model vs. that of independent case for different μ.

Table 2.10 tells us that when $\frac{1}{\mu}$ is changing around the CBO's maturity date ($T = 5$), then the longer the holding time of default effect, the larger the credit protection of the collateral pool. For example, as the holding times of default effect change from 1 year ($\mu = 1$) to 5.26 years ($\mu = 0.19$), the percentage changes of the credit protection with respect to the independent case increase from 26.415% to 35.352% when $b_2 = 2$.

Chapter 3

Pricing Defaultable Bonds with Regime Switching

3.1 Introduction

In chapter 2, we use indicator function to adjust one firm's intensity process if its correlated counterparter gets default. We also use an exponential distribution to control this indicator function. So this indicator function may be dropped from the firm intensity process in the future. It means that there is a possibility for the firm to get recovery after a certain period of time. This tells us that the intensity process shifts from one regime to another. In this chapter, we'll develop a generalized credit risk model which is subject to regime switching. It means that all the underlying factors in the credit risk model are subject to regime switching.

3.1.1 Review of Reduced Form Models

In [40], [13], they characterize default using intensity function which is assumed to be determinied by common economic factors as well as firm-specific factors. This achieves two effects. One is that the model can be applied to the situations where the underlying asset value is not observable. The other is that the default time is unpredictable, so this is consistent with the empirical literature that short-term debt often does not have zero credit spreads. The credit spread represents the premium that compensates the holder who bear the credit risk. The price or credit spread of a defaultable bond is directly related to a risk-free bond through default and recovery rates that both of them are defined exogenously. While the reduced-form

models have attractive properties, their main drawbacks are: the model is lack of a link between firm value and default; the credit spread generated by reduced-form model is still not large enough; and the model can not explain the jump part in credit spread. In order to explain the jump part in the credit spread, Arvanitis, Gregory and Laurent [2] directly model the credit spread. They used credit migration to generate jump part. So the model treats default as the consequence of credit migration rather than sudden occurrence. Jarrow and Yu [31] introduce a counterparty risk in their model. They assume that when one firm gets default, then its correlated firms may benefit or suffer from this default event. In order to achieve correlated default effect, they introduce an indicator function into hazard process.

There is an in-between approach which is developed by Cathcart and El-Jahel [5]. They provide a framework that combines structural and reduced-form approaches. By introducing a signaling process of uncertainty, they assume that a default event occurs in an expected or unexpected manner then the value of this signaling process reaches a certain lower barrier or at the first jump time of a hazard rate process. This signaling process of uncertainty represents the aggregation of all information on the quality of the firm currently available. The greater the value of the uncertain process, the poorer the quality of the firm. As Cathcart and El-Jahel [5], Schmid and Zagst [49] introduce this signaling process into their three factor model consisting of interest rate, credit spread and signaling process of uncertainty. From the practical point of view, their model is difficult to implement in practice.

Here, we present a model which can unify the current existing credit risk models by introducing regime shifts in the short interest rate process and hazard process. Our model can explain the following important issues: (1) credit spreads may change

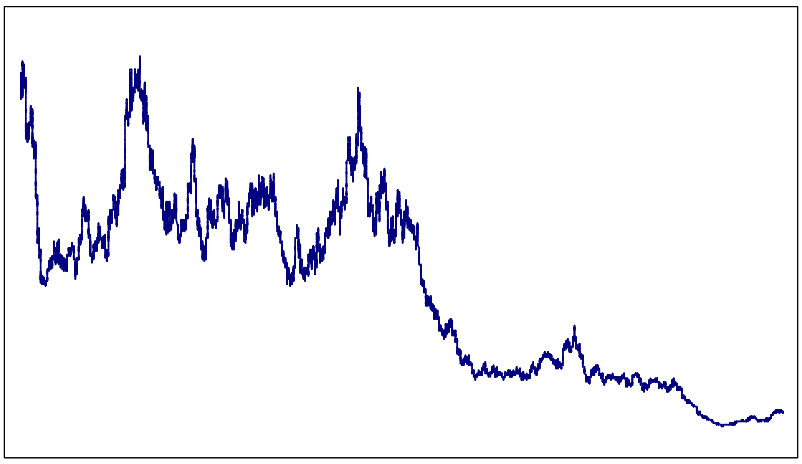

Figure 3.1: Time series with regime switching.

without default occurring; (2) credit spreads exhibit both a jump and a continuous component. (3) downward-sloping, upward-sloping, humped-sloping credit spread in short term can be generated by our model. Meanwhile, our model is much more tractable mathematically than the jump-diffusion type models or three factor models. And it is easy to implement in practice without losing any feature that the other models have. Some work on regime shifts has been done in literature. Zhang [59] introduces regime switching in stock liquidation. Yao, Zhang and Zhou [54] introduce regime switching in option pricing. Smith [50], and Bansal & Zhou [3] introduce regime switching into term structure of interest rate. All of them find sufficient evidences to support the regime switching models. Moreover, Smith (2002) also finds evidence that regime-switching model is favored over stochastic volatility model to represent the dynamic behavior of U.S. short-term interest rates.

3.2 Markov-Modulated Regime Switching

Many financial variables undergo episodes in which the behavior of the series seems to change quite dramatically. Graphically, it'll look like figure (3.1).

For the data plotted in figure (3.1), how should we model this process which is regime dependent. One way is that we can consider this process to be influenced by an unobserved random variable $\alpha(t)$, which will be called the regime that the process was in at time t. If $\alpha(t) = 1$, then the process is in regime 1, while $\alpha(t) = 2$ indicates that the process is in regime 2. For Markov-modulated regime switching, we assume that this unobserved random variable is a Markov chain.

3.3 Valuation of Defaultable Bonds

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a given probability space. All processes are assumed to be defined on this space and adapted to the filtration $(\mathcal{F}_t)$. The short-hand notation $E_t(\cdot)$ denotes $E(\cdot|\mathcal{F}_t)$, and all expectations are with respect to the measure $\mathbb{P}$. We work in an arbitrage-free setting and consider the behavior of the involved processes directly under an equivalent martingale measure $\mathbb{P}$.

Since $\mathbb{P}$ is an equivalent martingale measure, the money market account and default-free bond price are given by

$$B(t) = \exp\Big(\int_0^t r_s ds\Big)$$

and

$$p(t,T) = E_t\Big(\frac{B(t)}{B(T)}\Big)$$

, respectively, where r_t denotes the instantaneous default-free interest rate. In the default risk framework, a default appears at some random time τ. The payment of a defaultable bond consists of two parts:

(1) Given a maturity date $T > 0$, a random variable Y, which does not dependent on τ represents the promised payoffs - that is, the amount of cash the owner of the

claim will receive at time T, provided that the default has not occurred before the maturity date T.

(2) A predictable process X, pre-specified in the default-free world, models the payoff which is received if default occurs before maturity. The process is called the recovery process or the rebate.

So the price of the defaultable bond is, provided that the default has not occurred before time t,

$$V(t,T) = E_t\left(Y1_{\{\tau>T\}}\exp\left(-\int_t^T r_u du\right) + X_\tau 1_{\{\tau\leqslant T\}}\exp\left(-\int_t^\tau r_u du\right)\right)$$

In this paper, we follow Duffie and Singleton [13] and consider that default occurs at a rate of h_t. Here h_t is a given positive hazard rate process, i.e. the time-t hazard rate process, $h_t\delta_t$, gives the approximation probability of default for the bond over the time interval $(t, t+\delta_t)$. Suppose the promised payoff $Y = 1$, and use the recovery of market value, i.e. if default occurs, then the defaultable bond will be worth only a fraction of its predefault value, $X_\tau = \phi_\tau V(\tau-, T)$, where $0 \leq \phi_\tau < 1$. Under the equivalent martingale measure, using the result from Duffie and Singleton [13], the price of the defaultable bond is given by the expectation:

$$V(t,T,r,h) = E_t\left\{\exp(-\int_t^T R_u du)\right\}. \tag{3.1}$$

where $R_t = r_t + h_t L_t$ which is called adjusted discount rate, and L_t denotes the expected fractional loss in market value if default were to occur at time t. Comparing to the discount rate r_t in the default-free bond, this extra term $h_t L_t$ represents the " risk-neutral mean-loss rate". In this paper, we assume L_t is independent of t, i.e. $L_t = L$.

Duffee [11] uses an extented Kalman filter approach to test a square root diffusion model for the credit spread. He finds that the square root diffusion model is reasonably successful at fitting corporate bond yields. As in Duffee [11], we assume that the instantaneous interest rate r_t and hazard rate h_t follow mean reverting square-root processes (CIR model).

$$dr_t = a_1(t)(b_1(t) - r_t)dt + \sigma_1(t)\sqrt{r_t}dw_1(t). \tag{3.2}$$

and

$$dh_t = a_2(t)(b_2(t) - h_t)dt + \sigma_2(t)\sqrt{h_t}dw_2(t). \tag{3.3}$$

We assume that the correlation between dw_1 and dw_2 is ρ, i.e. $dw_1(t)dw_2(t) = \rho dt$.

Using Feynman-Kac formula, we know that $V(t) := V(t, T, r, h)$ must satisfy the following partial differential equation:

$$\begin{aligned} &\frac{\partial V}{\partial t} + a_1(b_1 - r_t)\frac{\partial V}{\partial r} + a_2(b_2 - h_t)\frac{\partial V}{\partial h} + \frac{1}{2}\sigma_1^2 V(t)\frac{\partial^2 V}{\partial r^2} + \frac{1}{2}\sigma_2^2 V(t)\frac{\partial^2 V}{\partial h^2} \\ &+ \rho\sigma_1\sigma_2\sqrt{r_t h_t}\frac{\partial^2 V}{\partial r \partial h} - (r_t + Lh_t)V(t) = 0, \end{aligned} \tag{3.4}$$

With boundary condition

$$V(T, T, r, h) = 1.$$

Our interest is to derive the implications for defaultable bond pricing when the interest rate and hazard rate are subject to regime shifts. Let $\{\alpha(t)\}$ denote a continuous-time Markov chain with state space $\mathcal{M} = \{1, 2, \cdots, m\}$. This finite-state Markov chain $\alpha(\cdot)$ can be used to represent the general market direction, the economy trend, etc. Let $Q = (q_{ij})_{m\times m}$ be the generator of $\alpha(t)$ with $q_{ij} \geq 0$ for $i \neq j$ and $\sum_{j=1}^{m} q_{ij} = 0$ for each $i \in \mathcal{M}$. Moreover, for any function f on $\mathcal{M}$, we denote $Qf(\cdot)(i) := \sum_{j=1}^{m} q_{ij}(f(j) - f(i))$. To keep things tractable, we will model the the

regime shifts process as a two-state Markov process ($m = 2$). Let the generator of $\alpha(\cdot)$ have the form

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}$$

with $\lambda_1 > 0$ and $\lambda_2 > 0$.

Using $\alpha(t)$, we have the following interest rate process and hazard rate process corresponding to equations (3.2) and (3.3).

$$dr_t = a_1(\alpha(t))(b_1(\alpha(t)) - r_t)dt + \sigma_1(\alpha(t))\sqrt{r_t}dw_1(t). \tag{3.5}$$

and

$$dh_t = a_2(\alpha(t))(b_2(\alpha(t)) - h_t)dt + \sigma_2(\alpha(t))\sqrt{h_t}dw_2(t). \tag{3.6}$$

Also using Feynman-Kac formula, we can verify that $V(i) := V(t, T, r, h, i)$ should satisfy the following system of PDE's:

$$\begin{aligned} &\frac{\partial V(i)}{\partial t} + a_1(i)(b_1(i) - r_t)\frac{\partial V(i)}{\partial r} + a_2(i)(b_2(i) - h_t)\frac{\partial V(i)}{\partial h} \\ &+ \frac{1}{2}\sigma_1^2(i)V(i)\frac{\partial^2 V(i)}{\partial r^2} + \frac{1}{2}\sigma_2^2(i)V(i)\frac{\partial^2 V(i)}{\partial h^2} + \rho\sigma_1(i)\sigma_2(i)\sqrt{r_t h_t}\frac{\partial^2 V(i)}{\partial r \partial h} \\ &- (r_t + Lh_t)V(i) + QV(t, T, r, h, \cdot)(i) = 0, i = 1, 2, \end{aligned} \tag{3.7}$$

with the boundary condition:

$$V(T, T, r, h, i) = 1, i = 1, 2.$$

Writing it separately, we get the following two partial differential equations:

$$\begin{aligned}
&\frac{\partial V(1)}{\partial t} + a_1(1)(b_1(1) - r_t)\frac{\partial V(1)}{\partial r} + a_2(1)(b_2(1) - h_t)\frac{\partial V(1)}{\partial h} \\
&+ \frac{1}{2}\sigma_1^2(1)V(1)\frac{\partial^2 V(1)}{\partial r^2} + \frac{1}{2}\sigma_2^2(1)V(1)\frac{\partial^2 V(1)}{\partial h^2} + \rho\sigma_1(1)\sigma_2(1)\sqrt{r_t h_t}\frac{\partial^2 V(1)}{\partial r \partial h} \\
&- (r_t + Lh_t)V(1) + \lambda_1(V(2) - V(1)) = 0.
\end{aligned} \tag{3.8}$$

And

$$\begin{aligned}
&\frac{\partial V(2)}{\partial t} + a_1(2)(b_1(2) - r_t)\frac{\partial V(2)}{\partial r} + a_2(2)(b_2(2) - h_t)\frac{\partial V(2)}{\partial h} \\
&+ \frac{1}{2}\sigma_1^2(2)V(2)\frac{\partial^2 V(2)}{\partial r^2} + \frac{1}{2}\sigma_2^2(2)V(2)\frac{\partial^2 V(2)}{\partial h^2} + \rho\sigma_1(2)\sigma_2(2)\sqrt{r_t h_t}\frac{\partial^2 V(2)}{\partial r \partial h} \\
&- (r_t + Lh_t)V(2) + \lambda_2(V(1) - V(2)) = 0.
\end{aligned} \tag{3.9}$$

with the boundary condition:

$$V(T, T, r, h, i) = 1, i = 1, 2.$$

where $V(i)$ denotes the value of defaultable bond when the Markov process is in state $i, i = 1, 2$.

Based on the above setup, analytical solution won't be easy to get. So we seek numerical approaches to calculate defaultable bond price.

3.4 Numerical Approaches

Finite difference approach, lattice (or tree) approach, and Monte Carlo simulation approach are the most popular vehicles for valuing derivative securities. When the number of underlying factors in the underlying system is less than three, finite difference approach and lattice approach are preferred to be used. However, when the number of underlying factors in your system is larger than three, Monte Carlo simulation will be easier to be implemented.

3.4.1 Explicit Finite Difference Method

The idea behind finite difference methods is to simplify the PDE by replacing the partial differentials with finite differences. There are three ways of implementing finite difference approach: explicit, implicit, and Crank-Nicolson. Each of them has its own advantages and disadvantages. Explicit finite difference method is very intuitive, and easy to implement comparing to implicit and Crank-Nicolson methods. The method's only disadvantage is that the numerical solution does not necessarily converge to the solution of the differential equation as the time step size δ_t tends to zero if we use the regular approximation method. Explicit finite difference method simply can be described as the unknown point can be calculated and expressed explicitly by known points. As we know, using explicit finite difference method to solve Black-Scholes PDE is equivalent to using trinomial tree model. By the trinomial tree model, we know that the coefficients before the three known points are the risk-neutral probabilities. So the three coefficients before the known points in the explicit finite difference methods serve the risk-neutral probabilities, therefore all three should be positive. However, we know that the sign of these coefficients also depend on the relative values of interest rate and stock volatility. Too small time and space steps will cause these coefficients to be negative. These will cause explicit finite difference method to be instable and lack of convergence. Hull & White [24] also pointed out that negative coefficients will happen when the underlying factor of PDE follows mean-reverting process.

However, there is a modification of explicit finite difference method introduced by Fleming & Soner [17]. They prove the convergence of explicit finite difference method by using viscosity solution approach. Here is the basic idea of the approximation method of Fleming & Soner [17]. Considering PDE (3.4), the value of a

defaultable bond V satisfies this partial differential equation. In this PDE, there are partial derivatives of V with respect to state variables r, h and time t. A small constant time interval, δ_t, a small constant change δ_r in r, and a small constant change δ_h in h are chosen. Then a grid is then constructed for considering values of V when r is equal to

$$r_0, r_0 + \delta_r, r_0 + 2\delta_r, \cdots, r_{\max},$$

,

$$h_0, h_0 + \delta_h, h_0 + 2\delta_h, \cdots, h_{\max},$$

and time is equal to

$$t, t + \delta_t, t + 2\delta_t, \cdots, T.$$

where the parameters r_0, h_0 and $r_{\max}, h_{\max}$ are the smallest and largest values of r, h, respectively, considered by the model, t is the current time, and T is the maturity date of the derivative security. Denote $t + n\delta_t$ by t_n, $r_0 + i\delta_r$ by r_i, $h_0 + j\delta_r$ by h_j, and the value of the derivative security at the (n, i, j) point on the grid by V_{ij}^n. The partial derivatives of V with respect to r at node (n, i, j) are approximated as follows,

if $a_1(b_1 - r_i) > 0$, then

$$\frac{\partial V}{\partial r} = \frac{V_{i+1,j}^n - V_{i,j}^n}{\delta_r}, \tag{3.10}$$

if $a_1(b_1 - r_i) < 0$, then

$$\frac{\partial V}{\partial r} = \frac{V_{i,j}^n - V_{i-1,j}^n}{\delta_r}. \tag{3.11}$$

Similarly, the partial derivatives of V with respect to h at node (n, i, j) are approximated as follows,

if $a_2(b_2 - h_j) > 0$, then

$$\frac{\partial V}{\partial h} = \frac{V_{i,j+1}^n - V_{i,j}^n}{\delta_h}, \tag{3.12}$$

if $a_2(b_2 - h_j) < 0$, then

$$\frac{\partial V}{\partial h} = \frac{V_{i,j}^n - V_{i,j-1}^n}{\delta_h}. \tag{3.13}$$

where $a_1(b_1 - r_i), a_2(b_2 - h_j)$ are the coefficients of partial derivatives V in PDE.

The second-order partial derivatives $\frac{\partial^2 V}{\partial r^2}$, $\frac{\partial^2 V}{\partial h^2}$ are approximated as follows,

$$\frac{\partial^2 V}{\partial r^2} = \frac{V^n_{i+1,j} - 2V^n_{ij} + V^n_{i-1,j}}{\delta_r^2} \tag{3.14}$$

and

$$\frac{\partial^2 V}{\partial h^2} = \frac{V^n_{i,j+1} - 2V^n_{ij} + V^n_{i,j-1}}{\delta_h^2} \tag{3.15}$$

The cross-term $\frac{\partial^2 V}{\partial r \partial h}$ is approximated as follows,

if $\rho > 0$, then

$$\frac{\partial^2 V}{\partial r \partial h} = \frac{(2V^n_{i,j} + V^n_{i+1,j+1} + V^n_{i-1,j-1}) - (V^n_{i+1,j} + V^n_{i-1,j} + V^n_{i,j+1} + V^n_{i,j-1})}{2\delta_r \delta_h} \tag{3.16}$$

if $\rho < 0$, then

$$\frac{\partial^2 V}{\partial r \partial h} = \frac{(V^n_{i+1,j} + V^n_{i-1,j} + V^n_{i,j+1} + V^n_{i,j-1}) - (2V^n_{i,j} + V^n_{i+1,j-1} + V^n_{i-1,j+1})}{2\delta_r \delta_h} \tag{3.17}$$

and the time derivative is approximated as

$$\frac{\partial V}{\partial t} = \frac{V^{n+1}_{i,j} - V^n_{i,j}}{\delta_t}. \tag{3.18}$$

According to the signs of coefficients of partial derivatives, we substitute corresponding approximation into PDE (3.4). Then we can calculate the value of $V^{n-1}_{i,j}$ at time t_{n-1} from the values of $V^n_{i,j}$ at time t_n. We have the following eight cases.

Case 1: if $a_1(b_1 - r_i) > 0, a_2(b_2 - h_j) > 0, \rho > 0$, then we have

$$\begin{aligned}
V_{i,j}^{n-1} &= (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j-1}^n \\
&+ (\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) - \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) - \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j+1}^n \\
&+ (\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n.
\end{aligned} \tag{3.19}$$

Case 2: if $a_1(b_1 - r_i) > 0, a_2(b_2 - h_j) > 0, \rho < 0$, then we have

$$\begin{aligned}
V_{i,j}^{n-1} &= (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j+1}^n \\
&+ (\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) - \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) - \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j-1}^n \\
&+ (\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n.
\end{aligned} \tag{3.20}$$

Case 3: if $a_1(b_1 - r_i) > 0, a_2(b_2 - h_j) < 0, \rho > 0$, then

$$\begin{aligned}
V_{i,j}^{n-1} &= (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j-1}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j+1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) - \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (-\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n \\
&+ (\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n.
\end{aligned} \tag{3.21}$$

Case 4: if $a_1(b_1 - r_i) > 0, a_2(b_2 - h_j) < 0, \rho < 0$, then

$$\begin{aligned}
V_{i,j}^{n-1} &= (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j+1}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j-1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) - \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (-\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n \\
&+ (\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n.
\end{aligned} \tag{3.22}$$

Case 5: if $a_1(b_1 - r_i) < 0, a_2(b_2 - h_j) > 0, \rho > 0$, then we have

$$\begin{aligned}
V_{i,j}^{n-1} &= (-\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n \\
&+ (\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) + \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) - \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j+1}^n \\
&+ (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j-1}^n.
\end{aligned} \tag{3.23}$$

Case 6: if $a_1(b_1 - r_i) < 0, a_2(b_2 - h_j) > 0, \rho < 0$, then we have

$$\begin{aligned}
V_{i,j}^{n-1} &= (-\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n \\
&+ (\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) + \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) - \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j-1}^n \\
&+ (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j+1}^n.
\end{aligned} \tag{3.24}$$

Case 7: if $a_1(b_1 - r_i) < 0, a_2(b_2 - h_j) < 0, \rho > 0$, then

$$\begin{aligned}
V_{i,j}^{n-1} &= (-\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j+1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) + \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (-\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n \\
&+ (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j-1}^n.
\end{aligned} \tag{3.25}$$

Case 8: if $a_1(b_1 - r_i) < 0, a_2(b_2 - h_j) < 0, \rho < 0$, then

$$\begin{aligned}
V_{i,j}^{n-1} &= (-\frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i-1,j}^n \\
&+ (\frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j+1}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i+1,j-1}^n \\
&+ (1 - \delta_t(r_i + Lh_j) + \frac{\delta_t}{\delta_r}a_1(b_1 - r_i) + \frac{\delta_t}{\delta_h}a_2(b_2 - h_j) \\
&- \frac{\delta_t}{\delta_r^2}\sigma_1^2 r_i - \frac{\delta_t}{\delta_h^2}\sigma_2^2 h_j - \frac{\delta_t}{\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j}^n \\
&+ (-\frac{\delta_t}{\delta_h}a_2(b_2 - h_j) + \frac{\delta_t}{2\delta_h^2}\sigma_2^2 h_j + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i,j-1}^n \\
&+ (\frac{\delta_t}{2\delta_r^2}\sigma_1^2 r_i + \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j})V_{i+1,j}^n - \frac{\delta_t}{2\delta_r\delta_h}\rho\sigma_1\sigma_2\sqrt{r_i h_j}V_{i-1,j+1}^n.
\end{aligned} \tag{3.26}$$

Since we know the value of V at time T, so the value of V at time t can be calculated by using (3.19)-(3.26) repeatedly to work back from the maturity date T to the current time t in step size of δ_t.

Now, we consider the boundary conditions. We have four points and four segments at each time t_n. Using Taylor expansion, we have

$$\begin{cases}
V(r_1) = V(r_0) + \frac{\partial V}{\partial r}\delta_r + \frac{1}{2}\frac{\partial^2 V}{\partial r^2}\delta_r^2 + o(\delta_r) \\
V(r_2) = V(r_0) + 2\frac{\partial V}{\partial r}\delta_r + \frac{1}{2}\frac{\partial^2 V}{\partial r^2}(2\delta_r)^2 + o(\delta_r)
\end{cases} \tag{3.27}$$

Solving the above equations, we have the approximations at lower boundary

$$\begin{aligned}\Big(\frac{\partial V}{\partial r}\Big)^n_{0,j} &= \frac{-V^n_{2,j} + 4V^n_{1,j} - 3V^n_{0,j}}{2\delta_r} \\ \Big(\frac{\partial^2 V}{\partial r^2}\Big)^n_{0,j} &= \frac{V^n_{2,j} - 2V^n_{1,j} + V^n_{0,j}}{(\delta_r)^2}.\end{aligned} \tag{3.28}$$

Similarly, the partial derivatives of V with respect to h at lower boundary can be approximated as follow,

$$\begin{aligned}\Big(\frac{\partial V}{\partial h}\Big)^n_{i,0} &= \frac{-V^n_{i,2} + 4V^n_{i,1} - 3V^n_{i,0}}{2\delta_h} \\ \Big(\frac{\partial^2 V}{\partial h^2}\Big)^n_{i,0} &= \frac{V^n_{i,2} - 2V^n_{i,1} + V^n_{i,0}}{(\delta_h)^2}.\end{aligned} \tag{3.29}$$

Using the same technique, we can get the following approximation at upper boundary

$$\begin{aligned}\Big(\frac{\partial V}{\partial r}\Big)^n_{N_r,j} &= \frac{V^n_{N_r-2,j} - 4V^n_{N_r-1,j} + 3V^n_{N_r,j}}{2\delta_r} \\ \Big(\frac{\partial^2 V}{\partial r^2}\Big)^n_{N_r,j} &= \frac{V^n_{N_r-2,j} - 2V^n_{N_r-1,j} + V^n_{N_r,j}}{(\delta_r)^2} \\ \Big(\frac{\partial V}{\partial h}\Big)^n_{i,N_h} &= \frac{V^n_{i,N_h-2} - 4V^n_{i,N_h-1} + 3V^n_{i,N_h}}{2\delta_h} \\ \Big(\frac{\partial^2 V}{\partial h^2}\Big)^n_{i,N_h} &= \frac{V^n_{i,N_h-2} - 2V^n_{i,N_h-1} + V^n_{i,N_h}}{(\delta_h)^2}.\end{aligned} \tag{3.30}$$

where $N_r\delta_r = r_{max}, N_h\delta_h = h_{max}$.

To use explicit finite difference approach on PDE system (3.7) with regime shifts, there is nothing new but to solve equations (3.8) and (3.9) simultaneously.

convergence: As Hull & White [24] pointed out, when using the explicit finite difference method and the underlying factor following a mean-reverting process, we'll have some problem on convergence which is caused by the mean-reverting process. One way to overcome this problem is to find the maximun/minimum values of the underlying factor that it could reach in specified time interval. They found an analytical solution for these maximum/minimum values. In this paper, we

use Monte Carlo simulation to find what the possible maximum/minimum values of interest rate and intensity rate could be in a given time interval.

3.4.2 Markov Chain Monte Carlo Simulation

Monte Carlo simulation has long been an important numerical tool for complex securities valuation problems. The technique is used extensively in the literature to obtain prices for instruments for which analytical solutions are not possible. Monte Carlo simulation provides a simple and flexible method for valuing these type of instruments. Here, our objective of using Monte Carlo simulation is to compare these results with those of explicit finite difference approach. Since there is Markov switching both in interest rate process and in hazard rate process, our simulation procedures are follows,

(1) Based on the given generator Q of this Markov chain, we generate a Markov chain in the span of given time-to-maturity of the defaultable bond.

(2) generate sample path of interest rate process, the coefficients in the drift and diffusion terms are decided by the state of the Markov chain (generated in step (1)) at that time.

(3) generate sample path of hazard rate process synchronously... as in step (2), however, this sample path is correlated to the one generated in step (2).

(4) use formula (1) to get the price of defaultable bond.

(5) repeat step (2) through (4) M times, then take the average of these M's defaultable bond prices.

Here, we use Markov chain Monte Carlo simulation as an alternative way to

calculate the defaultable bond price. Our objective is to check the stability of our results from explicit finite difference method.

3.5 Numerical Examples

Now we choose correlation $\rho = 0.5$, recovery rate $\delta = 0.4$, $\lambda = 10$, $\mu = 20$, and all coefficients in equations (3.5) and (3.6) as following

	$a_1(\alpha(t))$	$b_1(\alpha(t))$	$\sigma_1(\alpha(t))$	$a_2(\alpha(t))$	$b_2(\alpha(t))$	$\sigma_2(\alpha(t))$
$\alpha(t) = 1$	1.3	0.08	0.25	1.5	0.1	0.1
$\alpha(t) = 2$	2.5	0.05	0.25	2.5	0.05	0.25

Parameter table 3.1. Parameters for humped shape credit spread with $r_0 = 0.06, h_0 = 0.06$.

	$a_1(\alpha(t))$	$b_1(\alpha(t))$	$\sigma_1(\alpha(t))$	$a_2(\alpha(t))$	$b_2(\alpha(t))$	$\sigma_2(\alpha(t))$
$\alpha(t) = 1$	1.3	0.07	0.25	1.5	0.06	0.1
$\alpha(t) = 2$	2.5	0.03	0.15	2.5	0.15	0.25

Parameter table 3.2. Parameters for downward trend credit spread with $r_0 = 0.06, h_0 = 0.1$.

	$a_1(\alpha(t))$	$b_1(\alpha(t))$	$\sigma_1(\alpha(t))$	$a_2(\alpha(t))$	$b_2(\alpha(t))$	$\sigma_2(\alpha(t))$
$\alpha(t) = 1$	1.3	0.07	0.25	1.5	0.06	0.1
$\alpha(t) = 2$	1.3	0.07	0.25	2.5	0.15	0.25

Parameter table 3.3. Parameters for upward trend credit spread with $r_0 = 0.06, h_0 = 0.03$.

The default-free discount bond given by cox, Ingersoll, and Ross (1985) model:

$$p(l, r) = \exp(A(l) + B(l)r)$$

where

$$B(l) = \frac{-2[1 - \exp(-\gamma l)]}{2\gamma \exp(-\gamma l) + (a_1 - \gamma)[1 - \exp(-\gamma l)]}$$

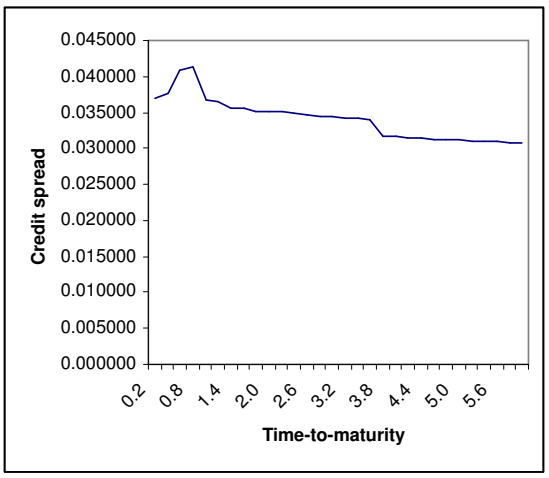

Figure 3.2: Humped shape credit spread

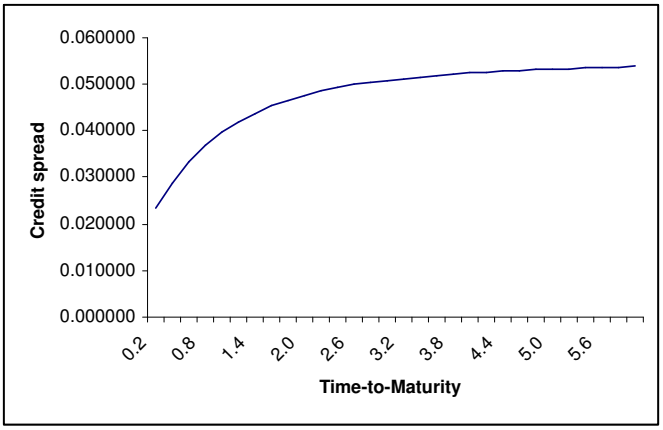

Figure 3.3: Upward trend spread.

$$\gamma := \sqrt{a_1^2 + 2\sigma_1^2}$$

and

$$A(l) = \frac{2a_1 b_1}{\sigma_1^2} \ln\left(\frac{2\gamma \exp[(a_1-\gamma)\frac{l}{2}]}{2\gamma \exp(-\gamma l) + (a_1-\gamma)[1-\exp(-\gamma l)]}\right)$$

where $l = T - t$

With downward trend credit spread, the firm improves their quality in the long run.

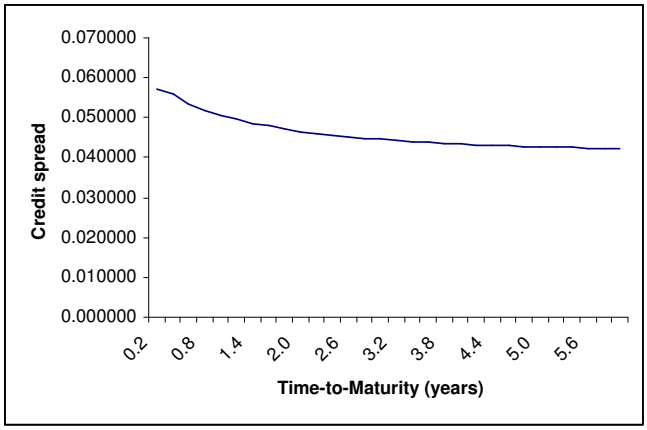

Figure 3.4: Downward trend spread.

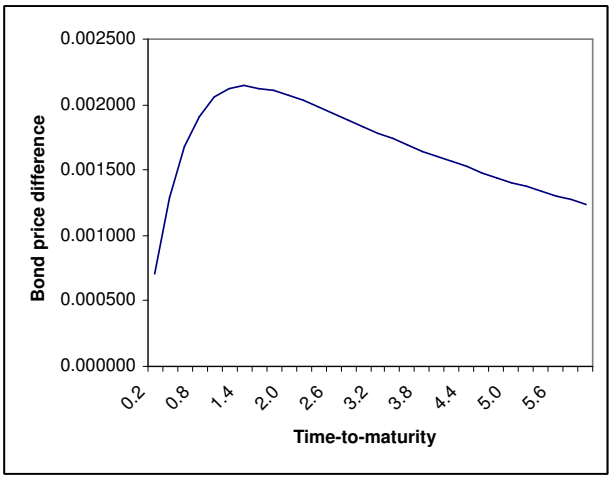

Figure 3.5: Impact of Markov switch on bond price.

Time-to-Maturity	Default-free	$\alpha(t) = 1$	$\alpha(t) = 2$	Spread
0.2	0.987604	0.980319	0.980796	0.037019
0.4	0.974595	0.960038	0.960859	0.037623
0.6	0.961188	0.937940	0.939123	0.040807
0.8	0.947542	0.916776	0.918161	0.041260
1.0	0.933773	0.900083	0.901331	0.036747
1.2	0.919968	0.880533	0.881821	0.036510
1.4	0.906191	0.862025	0.863296	0.035690
1.6	0.892487	0.843188	0.844458	0.035514
1.8	0.878891	0.824925	0.826180	0.035204
2.0	0.865427	0.806857	0.808097	0.035038
2.2	0.852114	0.788798	0.790012	0.035095
2.4	0.838963	0.771530	0.772721	0.034913
2.6	0.825984	0.754636	0.755803	0.034746
2.8	0.813182	0.738274	0.739413	0.034514
3.0	0.800561	0.722116	0.723230	0.034376
3.2	0.788122	0.706312	0.707403	0.034249
3.4	0.775866	0.690856	0.691923	0.034132
3.6	0.763794	0.675886	0.676926	0.033965
3.8	0.751903	0.666463	0.667425	0.031743
4.0	0.740194	0.652212	0.653152	0.031636
4.2	0.728663	0.638268	0.639188	0.031536
4.4	0.717310	0.624622	0.625522	0.031446
4.6	0.706131	0.611422	0.612301	0.031307
4.8	0.695126	0.598356	0.599216	0.031231
5.0	0.684291	0.585571	0.586413	0.031159
5.2	0.673624	0.573060	0.573883	0.031093
5.4	0.663123	0.560816	0.561621	0.031031
5.6	0.652785	0.548835	0.549624	0.030973
5.8	0.642608	0.537235	0.538005	0.030879
6.0	0.632590	0.525762	0.526515	0.030829

Result table 3.1. Humped shape credit spread.

Time-to-Maturity	Default-free	$\alpha(t) = 1$	$\alpha(t) = 2$	Spread	Price difference
0.2	0.987840	0.983199	0.982495	0.023545	0.000704
0.4	0.975453	0.964253	0.962959	0.028872	0.001294
0.6	0.962950	0.943923	0.942247	0.033262	0.001676
0.8	0.950408	0.922827	0.920920	0.036812	0.001907
1.0	0.937883	0.901414	0.899356	0.039660	0.002058
1.2	0.925414	0.879951	0.877834	0.041979	0.002117
1.4	0.913030	0.858672	0.856527	0.043844	0.002145
1.6	0.900752	0.837712	0.835586	0.045347	0.002126
1.8	0.888594	0.817031	0.814920	0.046646	0.002111
2.0	0.876567	0.796836	0.794763	0.047682	0.002073
2.2	0.864677	0.777106	0.775079	0.048536	0.002027
2.4	0.852929	0.757851	0.755874	0.049246	0.001977
2.6	0.841327	0.738937	0.737003	0.049911	0.001934
2.8	0.829872	0.720613	0.718732	0.050418	0.001881
3.0	0.818565	0.702808	0.700978	0.050823	0.001830
3.2	0.807406	0.685265	0.683480	0.051257	0.001785
3.4	0.796395	0.668156	0.666415	0.051639	0.001741
3.6	0.785530	0.651648	0.649957	0.051904	0.001691
3.8	0.774811	0.635626	0.633981	0.052108	0.001645
4.0	0.764236	0.619782	0.618178	0.052377	0.001604
4.2	0.753805	0.604333	0.602770	0.052621	0.001563
4.4	0.743514	0.589270	0.587745	0.052842	0.001525
4.6	0.733363	0.574795	0.573314	0.052963	0.001481
4.8	0.723350	0.560480	0.559035	0.053146	0.001445
5.0	0.713474	0.546785	0.545378	0.053218	0.001407
5.2	0.703731	0.533184	0.531812	0.053371	0.001372
5.4	0.694122	0.519922	0.518584	0.053513	0.001338
5.6	0.684644	0.506991	0.505687	0.053644	0.001304
5.8	0.675294	0.494378	0.493107	0.053767	0.001271
6.0	0.666073	0.482301	0.481065	0.053805	0.001236

Result table 3.2. Upward trend credit spread.

Time-to-Maturity	Default-free	$\alpha(t) = 1$	$\alpha(t) = 2$	Spread
0.2	0.987840	0.976603	0.976495	0.057202
0.4	0.975453	0.953928	0.953577	0.055786
0.6	0.962950	0.932628	0.932399	0.053326
0.8	0.950408	0.911770	0.911519	0.051880
1.0	0.937883	0.891604	0.891349	0.050603
1.2	0.925414	0.872010	0.871760	0.049534
1.4	0.913030	0.852893	0.852648	0.048668
1.6	0.900752	0.834186	0.833947	0.047983
1.8	0.888594	0.816060	0.815816	0.047307
2.0	0.876567	0.798586	0.798374	0.046585
2.2	0.864677	0.781316	0.781108	0.046080
2.4	0.852929	0.764396	0.764194	0.045663
2.6	0.841327	0.747987	0.747804	0.045229
2.8	0.829872	0.731825	0.731647	0.044904
3.0	0.818565	0.716072	0.715896	0.044591
3.2	0.807406	0.700605	0.700435	0.044338
3.4	0.796395	0.685485	0.685320	0.044108
3.6	0.785530	0.670688	0.670528	0.043904
3.8	0.774811	0.656541	0.656388	0.043588
4.0	0.764236	0.642393	0.642243	0.043419
4.2	0.753805	0.628730	0.628593	0.043198
4.4	0.743514	0.615196	0.615062	0.043056
4.6	0.733363	0.601956	0.601825	0.042925
4.8	0.723350	0.588997	0.588869	0.042807
5.0	0.713474	0.576374	0.576247	0.042678
5.2	0.703731	0.563971	0.563846	0.042576
5.4	0.694122	0.551834	0.551713	0.042482
5.6	0.684644	0.540151	0.540039	0.042330
5.8	0.675294	0.528534	0.528424	0.042249
6.0	0.666073	0.517169	0.517061	0.042172

Result table 3.3. Downward trend credit spread.

Time-to-Maturity	Finite difference	Monte Carlo	Price differences
0.2	0.976603	0.977697	-0.001094
0.4	0.953928	0.957697	-0.003769
0.6	0.932628	0.939085	-0.006457
0.8	0.911770	0.921496	-0.009726
1.0	0.891604	0.904520	-0.012916
1.2	0.872010	0.888074	-0.016064
1.4	0.852893	0.875112	-0.022219
1.6	0.834186	0.856403	-0.022217
1.8	0.816060	0.841080	-0.025020
2.0	0.798586	0.826144	-0.027558
2.2	0.781316	0.811368	-0.030052
2.4	0.764396	0.796986	-0.032590
2.6	0.747987	0.782771	-0.034784
2.8	0.731825	0.768880	-0.037055
3.0	0.716072	0.755143	-0.039071
3.2	0.700605	0.741794	-0.041189
3.4	0.685485	0.728568	-0.043083
3.6	0.670688	0.715548	-0.044860
3.8	0.656541	0.702977	-0.046436
4.0	0.642393	0.690361	-0.047968
4.2	0.628730	0.678108	-0.049378
4.4	0.615196	0.666098	-0.050902
4.6	0.601956	0.654206	-0.052250
4.8	0.588997	0.642554	-0.053557
5.0	0.576374	0.631174	-0.054800
5.2	0.563971	0.619917	-0.055946
5.4	0.551834	0.608837	-0.057003
5.6	0.540151	0.598144	-0.057993
5.8	0.528534	0.587383	-0.058849
6.0	0.517169	0.576994	-0.059825

Result table 3.4. Price from finite difference vs. price from Monte Carlo simulation.

For upward-sloping credit spread, it means that the possibility of default of the firm is getting bigger and bigger. The firm just could not keep its well performance, but deteriorate in quality over time.

For hump-shaped credit spread, it means that the firm has high possibility of default in the short and medium term, but the firm will improve its quality in the long term. So its credit spread widens in the short and medium term and tightens in the long term.

For downward-sloping credit spread, it means that the firm improves their quality as time goes.

As we can see from the above figures, our model can generate all kinds of credit spreads (downward trend, upward trend, and hump-shaped) with Markov regime switching. Comparing to Schmid & Zagst's [49] three-factor model (three-factor model is intractable in practice) and jump-diffusion model, our model is much easier to be implemented and more tractable. And the spread is not zero for short-term which is consistent with the real spread. It means that default could happen on a sudden.

3.6 Summary

In our paper, we incorporate regime shifts both in interest rate and in hazard rate. With these internal adjustments, our model can exhibit all the effects of credit spread model, counterparty risk model, three factor models, and jump-diffusion process model. Meanwhile, as Bansal and Zhou [3] point out, in order to account for the short interest rate data, incorporating regime shifts into the interest rate model is essential. Otherwise, multifactor version of CIR or affine models are needed. So we want to price accurate defaultable bond price, incorporating regime shifts into interest rate and hazard rate are crucial.

Chapter 4

Recursive Algorithms for Perpetual American Put Options

4.1 Introduction

Pricing American put options is equivalent to finding the stopping time when the put options reach their maximum values. It is well-known that some optimal stopping (or related free boundary problems) problems may be solved alternatively with probabilistic method. McKean [45] solved the optimal stopping point for perpetual American put options with no regime switching. In finance, many situations can be depicted by regime switching. For example, the dynamics of interest rate, exchange rate, stock price, etc. And regime switching is also widely used in finance, see [3], [21], [50], [54], [59]. Recently, Guo & Zhang [21] derived a closed-form solutions for perpetual American put options with regime switching. They consider a case when the Markov-modulated regime has only two states. However, if the underlying Markov chain has more than two states, a closed-form solution is difficult to obtain although the existence of solutions was proved in [59]. It is thus of practical interest to find feasible algorithms yielding good approximations to the optimal policy. With the motivation of reducing computational effort, a stochastic optimization procedure is developed.

4.2 Formulation

4.2.1 Hybrid Geometric Brownian Motion Model

Suppose that $\alpha(t)$ is a finite-state, continuous-time Markov chain with state space $\mathcal{M} = \{1, \ldots, m\}$, which represents market trends and other economic factors. For example, when $m = 2$, $\alpha(t) = 1$ stands for a bullish market, whereas $\alpha(t) = 2$ represents a bearish market. We may also consider, for instance, $\alpha(t) = (\alpha_1(t), \alpha_2(t))$, where $\alpha_1(t)$ models the market trends and $\alpha_2(t)$ represents the interest rates at time t. To take into account of more complex situation, we need to assume that the chain has more than two states, i.e., $m \geq 2$ in general. Let $S(t)$ be the price of the stock. We consider a hybrid geometric Brownian motion model that is risk neutral, in which $S(t)$ satisfies the stochastic differential equation

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu dt + \sigma(\alpha(t))dw(t), \\ S(0) &= S_0 \text{ initial price}, \end{aligned} \tag{4.1}$$

where $w(\cdot)$ is a real-valued standard Brownian motion that is independent of $\alpha(\cdot)$. The model is a hybrid geometric Brownian motion model (HGBM) or GBM with switching regime.

In (4.1), the volatility depends on the Markov chain $\alpha(t)$. Define another process

$$X(t) = \int_0^t r(\alpha(s))ds + \int_0^t \sigma(\alpha(s))dw(s), \tag{4.2}$$

where

$$r(i) = \mu - \frac{\sigma^2(i)}{2} \text{ for each } i = 1, \ldots, m. \tag{4.3}$$

Using $X(t)$, we can write the solution of (4.1) as

$$S(t) = S_0 \exp(X(t)). \tag{4.4}$$

Consider the perpetual American put options. The value functions take the form

$$v(S_0, i) = \sup_{\tau} E[\exp(-\mu\tau)(K - S(\tau)^+ \big| S(0) = S_0, \alpha(0) = i], \tag{4.5}$$

where τ is a stopping time to be specified shortly. For $\mathcal{M} = \{1, 2, \ldots, m\}$, a closed form solution has been found in [21]. Both dynamic programming approach and two-point-boundary-value method are used to approximate the solution of the optimal stopping problem. Note that due to the presence of the Markov chain, a system of value functions (a vector) value function must be dealt with. In what follows, we propose a stochastic approximation approach.

4.2.2 Method 1: Markov-dependent Procedure

Keeping in mind the threshold-type solutions, for $i \in \mathcal{M}$, let τ be a stopping time defined by

$$\tau = \inf\{t > 0 : (X(t), \alpha(t)) \notin D(x)\}, \tag{4.6}$$

where $x = (x^1, \ldots, x^m)$ with $x^1 \leq x^2 \leq \ldots \leq x^m$,

$$D(x) = \cup_{i=1}^{m}\{(x^i, \infty) \times \{i\}\}. \tag{4.7}$$

We aim at finding the optimal threshold level x_* so that the expected return is maximized. The problem can be rewritten as:

$$\text{Problem}\mathcal{P} : \begin{cases} \text{Find argmax } \varphi(x), \\ \varphi(x) = E[\exp(-\mu\tau)(K - S(\tau))^+], \end{cases} \tag{4.8}$$

where $\mu > 0$ is the discount rate. Use a stochastic optimization procedure to resolve the issue by constructing a sequence of estimates of the optimal threshold value x_* using

$$x_{n+1} = x_n + \{\text{step size}\} \cdot \{\text{gradient estimate of } \varphi(x_n)\},$$

where the step size is a decreasing sequence of real numbers or a small positive constant.

4.2.3 Gradient Estimates and Recursive Algorithm

The approximation procedures will depend on how the gradient estimates of $\varphi_x(x)$ are constructed. Let us begin with a simple noisy finite difference scheme. Several of its variants will be discussed in the subsequent sections. Using (4.1), generate a sample path of $X(t)$ that is the solution of (4.2). At time 0, choose initial estimate $x_0 = (x_0^1, \ldots, x_0^m)$. Compute τ_0 the first time that $(X(t), \alpha(t))$ reaches $(D(x))^c$, with

$$\tau_0 = \inf\{t \geq 0 : (X(t), \alpha(t)) \notin D(x_0)\}.$$

Choose the step size to be $\varepsilon_n = 1/n$, and let

$$t_n = \sum_{j=0}^{n-1} \varepsilon_j.$$

[Choose $\varepsilon_n = 1/(n+1)$ in the simulation.] Let

$$\xi_0 = (X(\tau_0), \tau_0),$$

and define the observable quantity

$$\widetilde{\varphi}(x_0, \xi_0) = \exp(-\mu\tau_0)(K - S(\tau_0))^+.$$

Then define the difference quotient

$$(D\widehat{\varphi}_0)^i = \frac{\widetilde{\varphi}(x_0^1, x_0^2, \ldots, x_0^i + \delta_0, x_0^{i+1}, \ldots, x_0^m, \xi_0^+) - \widetilde{\varphi}(x_0^1, x_0^2, \ldots, x_0^i - \delta_0, x_0^{i+1}, \ldots, x_0^m, \xi_0^-)}{2\delta_0},$$

where $(D\widetilde{\varphi}_0)^i$ denotes the ith component of the gradient estimate $D\widetilde{\varphi}_0$, and $\xi_0^{\pm}$ means that two different observations are used and δ_n is a sequence of real numbers satisfying $\delta_n \geq 0$ and $\delta_n \to 0$. [In the simulation, we can use $\delta_n = 1/(n+1)^{1/6}$.] Then compute $x_1 = (x_1^1, x_1^2, \ldots, x_1^m)$ according to

$$x_1^i = x_0^i + \varepsilon_0 (D\widehat{\varphi}_0)^i(x_0) I_{\{\alpha(\tau_0)=i\}}, \ i = 1, 2, \ldots, m.$$

Using induction, we then proceed construct the estimates recursively as follows. Suppose that $x_n = (x_n^1, \ldots, x_n^m)$ has been computed. Choose

$$\tau_n = \inf\{t : (X(t), \alpha(t)) \notin D(x_{n-1})\},$$

$$\xi_n = (X(\tau_n), \tau_n),$$

$$(D\widehat{\varphi}_n)^i = \frac{\widetilde{\varphi}(x_n^1, x_n^2, \ldots, x_n^i + \delta_n, x_n^{i+1}, \ldots, x_n^m, \xi_n^+) - \widetilde{\varphi}(x_n^1, x_n^2, \ldots, x_n^i - \delta_n, x_n^{i+1}, \ldots, x_n^m, \xi_n^-)}{2\delta_n}. \tag{4.9}$$

Then the stochastic approximation algorithm takes the form

$$x_{n+1}^i = x_n^i + \varepsilon_n (D\widehat{\varphi}_n)^i I_{\{\alpha(\tau_n)=i\}}, \ i = 1, 2, \ldots, m. \tag{4.10}$$

To ensure the boundedness of the iterates, use a projection algorithm

$$x_{n+1}^i = \Pi_{[\theta_l^i, \theta_u^i]}[x_n^i + \varepsilon_n (D\widehat{\varphi}_n)^i I_{\{\alpha(\tau_n)=i\}}], \quad \text{for} \quad i = 1, 2, \ldots, m, \tag{4.11}$$

where for each real value x,

$$\Pi_{[\theta_l^i, \theta_u^i]} x = \begin{cases} \theta_l^i, & \text{if } x < \theta_l^i, \\ \theta_u^i, & \text{if } x > \theta_u^i, \\ x, & \text{otherwise.} \end{cases}$$

The idea can be explained as follows. For component i, after the $x_n^i + \varepsilon_n (D\widehat{\varphi}_n)^i$ is computed, we compare its value with the bounds θ_l^i and θ_u^i. If the increment is smaller than the lower value θ_l^i, reset the value to θ_l^i, if it is larger than the upper value θ_u^i, reset its value to θ_u^i, otherwise keep its value as it was.

Convergence: the proof of convergence of this algorithm is similar to [57]. Several other stochastic recursive algorithms are also presented there.

4.3 Numerical Simulation

In this section, we consider a case with $m = 2$ and compare our approach with an analytical solution in [21]. We take

$$r = 3, \mu_1 = \mu_2 = 3, K = 5.$$

The simulation procedures are follows:

(1) For given Markov generator, say $\lambda_1 = 100, \lambda_2 = 100$, we use 0.0001 as time step-size to generate a Markov chain $\alpha(t)$.

(2) Based on this generated Markov chain $\alpha(t)$, we generate a sample path for $X(t)$. It means that the coefficients in expression $X(\cdot)$ at time t is determined by the generated markov chain. Then we use expression $S(t) = S_0 \exp(X(t))$ to get a sample path for stock price.

(3) Based on this generated sample path $S(t)$, use the proposed recursive algorithm to find the optimal threshold levels.

After 1000 iterations and averaging all threshold levels, we obtain the optimal threshold levels (x_1^*, x_2^*).

σ_1	7	8	9	10	11	12
Exact	(.646, .764)	(.531, .683)	(.441,.614)	(.369, .554)	(.312, .505)	(.266, .462)
Our results	(.549, .782)	(.505, .692)	(.468, .620)	(.427, .555)	(.394, .497)	(.357, .441)

Table 4.1. Dependency on σ_1 given $\sigma_2 = 5, \lambda_1 = \lambda_2 = 100$.

Keep all other parameters fixed, the threshold levels decrease as σ_1 increase. It implies high option premium, i.e. high put option value. Since the volatility of state

1 (σ_1) is larger than the volatility of state 2 ($\sigma_2 = 5$), the threshold level of state 1 is smaller than that of state 2.

λ_1	80	90	100	110	120	130
Exact	(.425, .596)	(.433, .605)	(.441,.614)	(.448, .621)	(.456, .629)	(.463, .637)
Our results	(.455, .592)	(.459, .615)	(.468, .620)	(.472, .625)	(.476, .634)	(.482, .649)

Table 4.2. Dependency on λ_1 given $\lambda_2 = 100, \sigma_1 = 9, \sigma_2 = 5$.

Keep all other parameters fixed, the threshold levels increase as λ_1 increases. The higher the λ_1 is, the shorter period the Markov chain stays in state 1. So these is a smaller weight on σ_1 which leads to a small average volatility, then a low option premium.

Bibliography

[1] Arvanitis, A.; and J. Gregory (2001) *Credit: The Complete Guide to Pricing, Hedging and Risk Management.* Risk books.

[2] Arvanitis, A.; Gregory, J.; and Laurent, J.-P. (1999) *Building Models for Credit Spreads.* The Journal of Derivatives, pp 27-43.

[3] Bansal, R.; and Zhou H. (2002) *Term Structure of Interest Rates with Regime Shifts.* Journal of Finance, Vol. LVII, no. 5, pp 1997-2043.

[4] Blanchet-Scalliet, C.; and Jeanblanc, M. (2001) *Hazard Rate for Credit Rate and Hedging Defaultable Contingent Calims.* preprint.

[5] Cathcart, L.; and El-Jahel L. (1998) *Valuation of Defaultable Bonds.* The Journal of Fixed Income, pp 65-78.

[6] Cheng, W.-y. (2001) *Recent Advances in Default Swap Valuation.* Journal of Derivatives, Fall, pp 18-27.

[7] Cox, J.C.; Ingersoll, J.E.; and Ross, S.A. (1985) *A Theory of the Term Structure of Interest Rates.* Econometrica, Vol. 53, No. 2, pp 385-407.

[8] Das, S. R.; and Tufano, P. (1996) *Pricing Credit-Sensitive Debt When Interest Rates, Credit Ratings and Credit Spreads are Stochastic.* Journal of Financial Engineering, Vol. 5, No. 2, pp 161-198.

[9] Davis, M.; and Lo, V. (2000) *Modelling Default Correlation in Bond Portfolios* preprint

[10] Duffee, G. R. (1998) *The Relation Between Treasury Yields and Corporate Bond Yield Spreads.* Journal of Finance, Vol. 53, No. 6, pp 2225-2241.

[11] Duffee, G. R. (1999) *Estimating the Price of Default Risk.* Review of Financial Studies, Vol. 12, No. 1, pp 197-226.

[12] Duffie, D. (1999) *Credit Swap Valuation.* Financial Analysts Journal, January/February, pp 73-87.

[13] Duffie, D.; and Singleton, K. J. (1999) *Modeling Term Structures of Defaultable Bonds.* Review of Financial Studies, Vol. 12, No. 4, pp 687-720.

[14] Duffie, D.; and Singleton, K. J. (2003) *Credit Risk: Pricing, Measurement, and Management.* Princeton series in Finance.

[15] Elliott, R. J.; Jeanblanc, M.; and Yor, M. (2000) *On Models of Default Risk.* Mathematical Finance, Vol. 10, No. 2, pp 179-195.

[16] Embrechts, P.; Lindskog, F.; and McNeil, A. (2001) *Modeling Dependence with Copula and Applications to Risk Management.* ETH preprint.

[17] Fleming, W. H.; and Soner, H. M. (1992) *Controlled Markov Prcesses and Viscosity Solutions.* Springer-Verlag, New York.

[18] Fons, J. S. (1994) *Using Default Rates to Model the Term Structure of Credit Risk* Financial Analysts Journal 50 (Sept./Oct.) pp 25-32

[19] Frees, E. W.; and Valdez, E. A. (1998) *Understanding Relationships Using Copulas.* North American Actuarial Journal, Vol. 2, No. 1, pp 1-25.

[20] Goodman, L.; (1998) *CBOs/CLOs: An Introduction.* Derivatives Quarterly, Fall, pp 21-27.

[21] Guo, X.; and Zhang, Q. (2002) *Closed-form Solutions for Perpetual American Put Options with Regime Switching.* preprint.

[22] Hamilton, J. D. (1994) *Time Series Analysis.* Princeton University Press.

[23] Hull, J. C.(1997) *Options, Futures, and Other Derivatives.* Prentice-Hall.

[24] Hull, J. C.; and White, A. (1990) *Valuing Derivative Securities Using the Explicit Finite Difference Method.* Journal of financial and Quantitative Analysis, Vol. 25, No. 1, p 87-100.

[25] Hull, J. C.; and White, A. (2000) *Valuing Credit Default Swaps I: No Counterparty Default Risk.* Journal of Fixed Income, Fall, pp 29-40.

[26] Hull, J. C.; and White, A. (2001) *Valuing Credit Default Swaps II: Modeling Default Correlations.* Journal of Fixed Income, Spring, pp 12-21.

[27] Jarrow, R. (2001) *Default Parameter Estimation Using Market Prices.* Financial Analysts Journal, September/October, pp 75-92.

[28] 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, Vol. 10, No. 2, pp 481-523.

[29] Jarrow, R. A.; and Turnbull, S. M. (1995) *Pricing Derivatives on Financial Securities Subject to Credit Risk.* Journal of Finance, Vol. 50, pp 53-85.

[30] Jarrow, R. A.; and Turnbull, S. M. (2000) *The Intersection of Market and Credit Risk.* Journal of Banking and Finance 24, pp 271-299.

[31] Jarrow, R.; and Yu, F. (2001) *Counterparty Risk and the Pricing of Defaultable Securities.* Journal of Finance, Vol. 56, No. 5, pp 1765-1799.

[32] Jeanblanc, M.; and Rutkowski, M. (1999) *Modeling of Default Risk: An Overview.* preprint.

[33] Kao, D.-L. (2000) *Estemating and Pricing Credit Risk: An Overview.* Association for investment Management and Research, July/August, pp 50-66.

[34] Karatzas, I.; and Shreve, S. E. (1998) *Methods of Mathematical Finance.* Springer, New York.

[35] Kijima, M. (2000) *Valuation of a Credit Swap of the Basket Type.* Review of Derivatives Research, 4, pp 81-97.

[36] Kijima, M.; and Muromacji, Y. (2000) *Credit Events and the Valuation of Credit Derivatives of Basket Type.* Revies of Derivatives Research, 4, pp 55-79.

[37] Kusuoka, S. (1999) *A Remark on Default Risk Models.* Advances in Mathematical Economics, 1, pp 69-82.

[38] Lando, D. (1994) *Three Essays on Contingent Claims Pricing.* Ph.D. Dissertation. Cornell University.

[39] Lando, D. (1997) *Modelling Bonds and Derivatives with default risk.* Cambridge Press, pp 369-393.

[40] Lando, D. (1998) *On Cox Processes and Credit Risky securities.* Review of Derivatives Research 2, pp 99-120.

[41] Lando, D. (2000) *Some Elements of Rating-Based Credit Risk Modeling.* Advanced Fixed-Income Valuation Tools, pp 193-215.

[42] Li, D. X. (2000) *On Default Correlation: A Copula Function Approach.* Journal of Fixed Income, Vol. 9, No. 4, pp 43-54.

[43] Longstaff, F.; and Schwartz, E. (1995) *A Simple Approach to Valuing Risky Fixed and Floating Rate Debt.* Journal of Finance, 50, 3, pp 789-819.

[44] Madan, D.; and Unal, H. (2000) *A Two-Factor Hazard Rate Model for Pricing Risky Debt and the Term Structure of Credit Spreads.* Journal of Financial and Quantitative Analysis, Vol. 35, No. 1, pp 43-65.

[45] McKean, H. P. (1965)*A Free Boundary Problem for the Heat Equation Arising from a Problem in Mathematical Economics*, Industrial Management Review, 6 pp 32–39.

[46] Merton, R. C. (1974) *On the Pricing of Corporate Debt: The Risk Structure of Interest Rates.* Journal of Finance, 29, pp 449-470.

[47] Musiela, M.; and Rutkowski, M. (1997) . *Martingale Methods in Financial Modelling.* Springer-Verlag, New York.

[48] Nelsen, R. B. (1999) *An Introduction to Copulas.* Springer, New York.

[49] Schmid, B.; and Zagst, R. (2000) *A Three-Factor Defaultable Term Structure Model.* The Journal of Fixed Income, pp 63-79.

[50] Smith, D. R. (2002) *Markov-Switching and Stochastic Volatility Diffusion Models of Short-Term Interest Rates.* Journal of Business & Economic Statistics, Vol. 20, No. 2, pp 183-197.

[51] Skora, R. K. (1998) *Correlation-the Hidden Risk in Collateralized Debt Obligations.* Skora & Company Inc.

[52] Wong, M. C. W.; and Hodges S. D. *Pricing Defaultable Coupon Bonds Under a Jump-Diffusion Process.* The Journal of Fixed Income, p 51-64, June 2002.

[53] Wong, P.-S. (1998) *A Unifying Credit Model.* Ph.D. Dissertation. Carnegie Mellon University.

[54] Yao, D.; Zhang Q.; and Zhou X. Y. (2001) *Option Pricing with Markov-Modulated Volatility.* prepeint.

[55] Yin, G. (1999) *Rates of Convergence for a Class of Global Stochastic Optimization Algorithms.* SIAM J. Optim., 10, pp99-120.

[56] Yin, G.; and Zhang, Q. (1998) *Continuous-Time Markov Chains and Application: a Singular Perturbation Approach.* Springer-Verlag, New York, 1998.

[57] Yin, G.; Liu, R.; and Zhang, Q. (2002) *Recursive Algorithms for Stock Liquidation: A Stochastic Optimization Approach*, preprint.

[58] Yu, F. (1999) *Three Essays on the Pricing and Management of Credit Risk.* Ph.D. Dissertation. Cornell University.

[59] Zhang, Q. (2001) *Stock Trading: an Optimal Selling Rule.* SIAM J. of Control. Optim. 40, pp64-87.

[60] Zhou, C. (1997) *A Jump Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities.* Federal Reserve Board, Washington.

Appendix A

Derivation of Equation (2.8)

$$\mathbb{P}(\tau_B > s \big| \mathcal{G}_t^B) = \mathbb{P}(\tau_B > s \big| \mathcal{G}_t \vee \mathcal{D}_t^A)$$

$$= 1_{\{\tau_A > t\}} \frac{\mathbb{P}(\tau_B > s, \tau_A > t \big| \mathcal{G}_t)}{\mathbb{P}(\tau_A > t \big| \mathcal{G}_t)} + 1_{\{\tau_A \leqslant t\}} \frac{\mathbb{P}(\tau_B > s, \tau_A \leqslant t \big| \mathcal{G}_t)}{\mathbb{P}(\tau_A \leqslant t \big| \mathcal{G}_t)}$$

$$= 1_{\{\tau_A > t\}} \frac{\mathbb{P}(\tau_B > s, \tau_A > t \big| \mathcal{G}_t)}{\mathbb{P}(\tau_A > t \big| \mathcal{G}_t)} +$$

$$+ 1_{\{\tau_A \leqslant t\}} \frac{\mathbb{P}(\tau_B > s \big| \mathcal{G}_t) - \mathbb{P}(\tau_B > s, \tau_A > t \big| \mathcal{G}_t)}{1 - \mathbb{P}(\tau_A > t \big| \mathcal{G}_t)} \tag{A.1}$$

$$= \begin{cases} \frac{E(\mathbb{P}(\tau_B > s, \tau_A > t | \mathcal{G}_{T^*}) | \mathcal{G}_t)}{e^{-\Gamma_t^A}}, & if \quad 1_{\{\tau_A > t\}} = 1 \\ \\ \frac{E(e^{-\Gamma_s^B} \big| \mathcal{G}_t) - E(\mathbb{P}(\tau_B > s, \tau_A > t | \mathcal{G}_{T^*}) | \mathcal{G}_t)}{1 - e^{-\Gamma_t^A}}, & if \quad 1_{\{\tau_A \leqslant t\}} = 1. \end{cases}$$

$$= \begin{cases} \frac{E(\mathbb{P}(E_1 < e^{-\Gamma_s^B}, E_2 < e^{-\Gamma_t^A} \big| \mathcal{G}_{T^*}) | \mathcal{G}_t)}{e^{-\Gamma_t^A}}, & if \quad 1_{\{\tau_A > t\}} = 1 \\ \\ \frac{E(e^{-\Gamma_s^B} \big| \mathcal{G}_t) - E(\mathbb{P}(E_1 < e^{-\Gamma_s^B}, E_2 < e^{-\Gamma_t^A} \big| \mathcal{G}_{T^*}) | \mathcal{G}_t)}{1 - e^{-\Gamma_t^A}}, & if \quad 1_{\{\tau_A \leqslant t\}} = 1. \end{cases}$$

$$= \begin{cases} e^{\Gamma_t^A} E[C(e^{-\Gamma_s^B}, e^{-\Gamma_t^A})\big|\mathcal{G}_t], & if \quad 1_{\{\tau_A > t\}} = 1 \\ \\ \frac{E\left(e^{-\Gamma_s^B}\big|\mathcal{G}_t\right) - E[C(e^{-\Gamma_s^B}, e^{-\Gamma_t^A})\big|\mathcal{G}_t]}{1 - e^{-\Gamma_t^A}}, & if \quad 1_{\{\tau_A \leqslant t\}} = 1. \end{cases} \tag{A.2}$$

Appendix B

Derivation of Equations (2.11) and (2.12)

We use u_1, u_2, v to denote $e^{-\int_0^s \lambda_B u du}, e^{-\int_0^t \lambda_B u du}, e^{-\int_0^t \lambda_A u du}$, respectively. Using Farlie-Gumbel-Morgenstern copula $C(u,v) = uv(1+\alpha(1-u)(1-v))$, we have

$$\mathbb{P}(\tau_B > s|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = 1_{\{\tau_B > t, \tau_A \leqslant t\}} \frac{\mathbb{P}(\tau_B > s, \tau_A \leqslant t|\mathcal{G}_t)}{\mathbb{P}(\tau_B > t, \tau_A \leqslant t|\mathcal{G}_t)}$$

$$\begin{aligned} &= \frac{E(\mathbb{P}(\tau_B > s|\mathcal{G}_{T^*}) - \mathbb{P}(\tau_B > s, \tau_A > t|\mathcal{G}_{T^*})|\mathcal{G}_t)}{E(\mathbb{P}(\tau_B > t|\mathcal{G}_{T^*}) - \mathbb{P}(\tau_B > t, \tau_A > t|\mathcal{G}_{T^*})|\mathcal{G}_t)} \\ &= \frac{E(u_1 - u_1 v(1+\alpha(1-u_1)(1-v))|\mathcal{G}_t)}{E(u_2 - u_2 v(1+\alpha(1-u_2)(1-v))|\mathcal{G}_t)} \\ &= \frac{E(u_1(1-v)[1-\alpha v(1-u_1)]|\mathcal{G}_t)}{E(u_2(1-v)[1-\alpha v(1-u_2)]|\mathcal{G}_t)} \end{aligned} \tag{B.1}$$

Since u_2, v are $\mathcal{G}_t$-measurable, so $E(u_2(1-v)[1-\alpha v(1-u_2)]|\mathcal{G}_t) = u_2(1-v)[1-\alpha v(1-u_2)]$. Therefore, the above equation can be simplified as $E(e^{-\int_t^s \lambda_B(u)du}\frac{1-\alpha v(1-u_1)}{1-\alpha v(1-u_2)}|\mathcal{G}_t)$. Now we use approximation formula $e^x \doteq 1+x$, so $1-\alpha v(1-u_1) \doteq e^{-\alpha v(1-u_1)}, 1-\alpha v(1-u_2) \doteq e^{-\alpha v(1-u_2)}$, and $1-u_1 \doteq \int_0^s \lambda_B(u)du, 1-u_2 \doteq \int_0^t \lambda_B(u)du$. so

$$E(e^{-\int_t^s \lambda_B(u)du}\frac{1-\alpha v(1-u_1)}{1-\alpha v(1-u_2)}|\mathcal{G}_t) = E(e^{-\int_t^s (1+\alpha\int_0^t \lambda_A(u)du)\lambda_B(u)du}|\mathcal{G}_t) \tag{B.2}$$

This is the proof of equation (2.11). Similarly, we can prove equation (2.12).

Appendix C

Derivation of Equation (2.14)

$$
\begin{aligned}
&\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = \mathbb{E}(\exp(-\int_t^T \lambda_B(u)du)|\mathcal{G}_t) \\
&= \mathbb{E}(\exp(-\int_t^T [b_1 + b_2 1_{\{u \leqslant S+\eta\}}]ds)|\mathcal{G}_t) \\
&= e^{-b_1(T-t)}\mathbb{E}(\exp(-b_2\int_t^T 1_{\{u \leqslant S+\eta\}}|\mathcal{G}_t)) \\
&= e^{-b_1(T-t)}\left\{\int_0^{t-S} \mu e^{-\mu y}dy + \int_{t-S}^{T-S} e^{-b_2(S+y-t)}\mu e^{-\mu y}dy + \int_{T-S}^{\infty} e^{-b_2(T-t)}\mu e^{-\mu y}dy\right\} \\
&= e^{-b_1(T-t)}\left\{1 - \frac{b_2}{b_2+\mu}e^{-\mu(t-S)} + \frac{b_2}{b_2+\mu}e^{-b_2(T-t)-\mu(T-S)}\right\}
\end{aligned}
\tag{C.1}
$$

Appendix D

Derivation of Equations (2.17), (2.18) and (2.19)

From (2.12), it's sufficient to derive $E_t[\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}} ds)]$. We know that $\mathbb{P}(\tau_A > s|\mathcal{G}_t) = e^{-a(s-t)}$. We consider two cases: $\eta = y > T-t$ and $\eta = y \leqslant T-t$.

When $\eta = y > T-t$, then $\tau_A + \eta|_{\eta=y} > T$, So

$$\int_{T-t}^{\infty} E_t(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A+y\}} ds)|\eta = y)\mu e^{-\mu y} dy$$

$$= \int_{T-t}^{\infty} E_t(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s\}} ds))\mu e^{-\mu y} dy \tag{D.1}$$

$$= \int_{T-t}^{\infty} E_t(e^{-b_2(T-\tau_A)1_{\{\tau_A \leqslant T\}}})\mu e^{-\mu y} dy$$

From Jarrow & Yu [31], we know that

$$E_t(e^{-b_2(T-\tau_A)1_{\{\tau_A \leqslant T\}}}) = \begin{cases} (a(T-t)+1)e^{-a(T-t)}, & if \quad a = b_2 \\ \\ \frac{b_2 e^{-a(T-t)} - a e^{-b_2(T-t)}}{b_2 - a}, & if \quad a \neq b_2. \end{cases}$$

Therefore, we have

$$\int_{T-t}^{\infty} E_t(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + y\}} ds) | \eta = y) \mu e^{-\mu y} dy$$

$$= \begin{cases} \int_{T-t}^{\infty} (a(T-t)+1) e^{-a(T-t)} \mu e^{-\mu y} dy, & if \quad a = b_2 \\ \\ \int_{T-t}^{\infty} \frac{b_2 e^{-a(T-t)} - a e^{-b_2(T-t)}}{b_2 - a} \mu e^{-\mu y} dy, & if \quad a \neq b_2. \end{cases} \tag{D.2}$$

$$= \begin{cases} e^{-\mu(T-t)} (a(T-t)+1) e^{-a(T-t)}, & if \quad a = b_2 \\ \\ e^{-\mu(T-t)} \frac{b_2 e^{-a(T-t)} - a e^{-b_2(T-t)}}{b_2 - a}, & if \quad a \neq b_2. \end{cases}$$

When $\eta = y \leqslant T - t$, we have

$$E_t[\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + y\}} ds) | \eta = y] \tag{D.3}$$

$$= \int_t^{T-y} e^{-b_2 y} a e^{-a(x-t)} dx + \int_{T-y}^{T} e^{-b_2(T-x)} a e^{-a(x-t)} dx + \int_T^{\infty} a e^{-a(x-t)} dx$$

Now we're going to simplify the above three integrals, then take the integration from 0 to $T - t$ w.r.t. η. Let

$$I_1 = \int_t^{T-y} e^{-b_2 y} a e^{-a(x-t)} dx = e^{-b_2 y} - e^{-a(T-t)+(a-b_2)y}$$

Then

$$\int_0^{T-t} I_1 \mu e^{-\mu y} dy = \int_0^{T-t} \mu e^{-(b_2+\mu)y} dy - \int_0^{T-t} \mu e^{-a(T-t)} e^{(a-b_2-\mu)y} dy$$

$$= \begin{cases} \frac{\mu}{b_2+\mu}[1 - e^{-(b_2+\mu)(T-t)}] - \mu(T-t)e^{-a(T-t)}, & if \quad a = b_2 + \mu \\ \\ \frac{\mu}{b_2+\mu}[1 - e^{-(b_2+\mu)(T-t)}] + \frac{\mu e^{-a(T-t)}}{a-b_2-\mu}[1 - e^{(a-b_2-\mu)(T-t)}], & if \quad a \neq b_2 + \mu. \end{cases} \tag{D.4}$$

Let

$$I_2 = \int_{T-y}^{T} e^{-b_2(T-x)} a e^{-a(x-t)} dx = \begin{cases} aye^{-b_2 T + at}, & if \quad a = b_2 \\ \\ \frac{ae^{-a(T-t)}}{b_2-a}[1 - e^{(a-b_2)y}], & if \quad a \neq b_2. \end{cases}$$

Then, we have:

When $a = b_2$

$$\int_0^{T-t} I_2 \mu e^{-\mu y} dy = \int_0^{T-t} aye^{-b_2 T + at} \mu e^{-\mu y} dy \tag{D.5}$$

$$= ae^{-b_2 T + at}[\frac{1 - e^{-\mu(T-t)}}{\mu} - (T-t)e^{-\mu(T-t)}]$$

When $a \neq b_2$

$$\int_0^{T-t} I_2 \mu e^{-\mu y} dy = \int_0^{T-t} \frac{ae^{-a(T-t)}}{b_2 - a}[1 - e^{(a-b_2)y}] \mu e^{-\mu y} dy$$

$$= \begin{cases} \frac{ae^{-a(T-t)}}{b_2-a}[1 - e^{-\mu(T-t)} - \mu(T-t)], & if \quad a = b_2 + \mu \\ \\ \frac{ae^{-a(T-t)}}{b_2-a}[\frac{a-b_2}{a-b_2-\mu} - e^{-\mu(T-t)} - \frac{\mu}{a-b_2-\mu}e^{(a-b_2-\mu)(T-t)}], & if \quad a \neq b_2 + \mu. \end{cases} \tag{D.6}$$

Let

$$I_3 = \int_T^{\infty} a e^{-a(x-t)} dx = e^{-a(T-t)}.$$

So

$$\int_0^{T-t} I_3 \mu e^{-\mu y} dy = \int_0^{T-t} e^{-a(T-t)} \mu e^{-\mu y} dy = e^{-a(T-t)}(1 - e^{-\mu(T-t)}). \tag{D.7}$$

Using equations (D.5), (D.6) and (D.7), we can get $E_t[\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}} ds)]$. Therefore, we have following equations:

When $a \neq b_2, a \neq b_2 + \mu$, then

$$E(\exp(-\int_t^T \lambda_B(s) ds)|\mathcal{G}_t) = e^{-b_1(T-t)} \Bigg\{ \{e^{-a(T-t)} - e^{-(a+\mu)(T-t)}\}$$

$$+ \{\frac{\mu}{b_2+\mu}[1 - e^{-(b_2+\mu)(T-t)}] + \frac{\mu e^{-a(T-t)}}{a-b_2-\mu}[1 - e^{(a-b_2-\mu)(T-t)}]\}$$

$$+ \{\frac{a e^{-a(T-t)}}{b_2 - a}[\frac{a-b_2}{a-b_2-\mu} - e^{-\mu(T-t)} - \frac{\mu}{a-b_2-\mu} e^{(a-b_2-\mu)(T-t)}]\}$$

$$+ e^{-\mu(T-t)} \frac{b_2 e^{-a(T-t)} - a e^{-b_2(T-t)}}{b_2 - a} \Bigg\}$$

$$= e^{-b_1(T-t)} \Bigg\{ \frac{\mu}{b_2+\mu} - \frac{b_2}{a-b_2-\mu} e^{-a(T-t)} + \frac{a b_2}{(b_2+\mu)(a-b_2-\mu)} e^{-(b_2+\mu)(T-t)} \Bigg\}. \tag{D.8}$$

When $a = b_2, a \neq b_2 + \mu$, then

$$E(\exp(-\int_t^T \lambda_B(s) ds)|\mathcal{G}_t) \tag{D.9}$$

$$= e^{-b_1(T-t)} \{\frac{\mu}{a+\mu} + \frac{a}{\mu} e^{-a(T-t)} - \frac{a^2}{\mu(\mu+a)} e^{-(a+\mu)(T-t)}\}$$

When $a \neq b_2, a = b_2 + \mu$, then

$$E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t) = e^{-b_1(T-t)}\Bigg\{\{e^{-a(T-t)} - e^{-(a+\mu)(T-t)}\}$$

$$+ \{\frac{\mu}{b_2+\mu}[1 - e^{-(b_2+\mu)(T-t)}] - \mu(T-t)e^{-a(T-t)}\}$$

$$+ \{\frac{ae^{-a(T-t)}}{b_2-a}[1 - e^{-\mu(T-t)} - \mu(T-t)]\}$$

$$+ e^{-\mu(T-t)}\frac{b_2e^{-a(T-t)} - ae^{-b_2(T-t)}}{b_2-a}\Bigg\}$$

$$= e^{-b_1(T-t)}\Bigg\{\frac{\mu}{b_2+\mu} + \frac{b_2(1-\mu(T-t))}{b_2-a}e^{-a(T-t)} - \frac{b_2(\mu+a)}{(b_2-a)(b_2-\mu)}e^{-(b_2+\mu)(T-t)}\Bigg\}. \tag{D.10}$$

Appendix E

Derivation of Equation (2.22)

We know that

$$\mathbb{P}(\tau_A > s|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = \exp(-\int_t^s \lambda_A(u)du) = e^{-a_0(s-t)-a_1 R_{t,s}}.$$

Where $R_{t,s} = \int_t^s r(u)du$.

We let $E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}}))$ denote $E(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}})|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)$. Using $R_{t,T}$ is $\mathcal{G}_{T^*}$-measurable, we have

$$\mathbb{P}(\tau_B > T|\mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(s)ds)|\mathcal{G}_t)$$

$$= E(\exp(-\int_t^T [b_0 + b_1 r(s) + b_2 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}}]ds)|\mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(E(\exp(-\int_t^T [b_1 r(s) + b_2 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}}]ds)|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)|\mathcal{G}_t)$$

$$= ^{-b_0(T-t)} E(\exp(-b_1 \int_t^T r(s)ds) E(\exp(-\int_t^T b_2 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}}ds)|\mathcal{G}_{T^*} \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B)|\mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(\exp(-b_1 \int_t^T r(s)ds) E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A+\eta\}}ds))|\mathcal{G}_t) \tag{E.1}$$

Now we use the property that $E(X) = E(E(X|Y))$, where X, Y are random variables. As we did in appendix D, we consider two cases: $\eta = y > T - t$ and $\eta = y \leqslant T - t$.

When $\eta = y > T - t$, then $\tau_A + \eta|_{\eta=y} > T$, So

$$E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}} ds))$$

$$= \int_{T-t}^{\infty} E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + y\}} ds) | \eta = y) \mu e^{-\mu y} dy$$

$$= \int_{T-t}^{\infty} E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s\}} ds)) \mu e^{-\mu y} dy \tag{E.2}$$

$$= \int_{T-t}^{\infty} E_{t,T^*}(e^{-b_2(T-\tau_A)1_{\{\tau_A \leqslant T\}}}) \mu e^{-\mu y} dy$$

And

$$E_{t,T^*}(e^{-b_2(T-\tau_A)1_{\{\tau_A \leqslant T\}}}) = \int_t^T e^{-b_2(T-x)} d(1 - e^{-a_0(x-t) - a_1 R_{t,x}})$$

$$+ \int_T^{\infty} d(1 - e^{-a_0(x-t) - a_1 R_{t,x}})$$

$$= e^{-b_2(T-x)}(1 - e^{-a_0(x-t) - a_1 R_{t,x}})\Big|_t^T - \int_t^T (1 - e^{-a_0(x-t) - a_1 R_{t,x}}) b_2 e^{-b_2(T-x)} dx \tag{E.3}$$

$$+ (1 - e^{-a_0(x-t) - a_1 R_{t,x}})\Big|_T^{\infty}$$

$$= e^{-b_2(T-t)}(1 + b_2 \int_t^T e^{-(a_0 - b_2)(x-t) - a_1 R_{t,x}} dx).$$

Therefore, for $\eta = y > T - t$, we have:

$$E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}} ds))$$

$$= \int_{T-t}^{\infty} e^{-b_2(T-t)}(1 + b_2 \int_t^T e^{-(a_0-b_2)(x-t)-a_1 R_{t,x}} dx)\mu e^{-\mu y} dy \tag{E.4}$$

$$= e^{-(b_2+\mu)(T-t)}(1 + b_2 \int_t^T e^{-(a_0-b_2)(x-t)-a_1 R_{t,x}} dx).$$

When $\eta = y \leqslant T - t$, we have:

$$E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}} ds)) \tag{E.5}$$

$$= \int_0^{T-t} E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}})|_{\eta=y})\mu e^{-\mu y} dy.$$

And

$$E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + y\}} ds)) = \int_t^{T-y} e^{-b_2 y} d(1 - e^{-a_0(x-t)-a_1 R_{t,x}})$$

$$+ \int_{T-y}^{T} e^{-b_2(T-x)} d(1 - e^{-a_0(x-t)-a_1 R_{t,x}}) + \int_T^{\infty} d(1 - e^{-a_0(x-t)-a_1 R_{t,x}})$$

$$= e^{-b_2 y}(1 - e^{-a_0(x-t)-a_1 R_{t,x}})\Big|_t^{T-y} + e^{-b_2(T-x)}(1 - e^{-a_0(x-t)-a_1 R_{t,x}})\Big|_{T-y}^{T} \tag{E.6}$$

$$- b_2 \int_{T-y}^{T} (1 - e^{-a_0(x-t)-a_1 R_{t,x}}) e^{-b_2(T-x)} dx + (1 - e^{-a_0(x-t)-a_1 R_{t,x}})\Big|_T^{\infty}$$

$$= e^{-b_2 y} + b_2 e^{-b_2(T-t)} \int_{T-y}^{T} e^{(b_2-a_0)(x-t)-a_1 R_{t,x}} dx.$$

Therefore, for $\eta = y \leqslant T - t$, we have:

$$E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}} ds))$$

$$= \int_0^{T-t} E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}})|_{\eta=y}) \mu e^{-\mu y} dy \tag{E.7}$$

$$= \frac{\mu}{\mu + b_2} + \mu b_2 e^{-b_2 T + a_0 t} \int_0^{T-t} \int_{T-y}^T [e^{(b_2 - a_0)x - a_1 R_{t,x}} dx] e^{-\mu y} dy.$$

Now combining equations (E.3) and (E.6), we have:

$$\mathbb{P}(\tau_B > T | \mathcal{G}_t \vee \mathcal{D}_t^A \vee \mathcal{D}_t^B) = E(\exp(-\int_t^T \lambda_B(s) ds) | \mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(\exp(-b_1 \int_t^T r(s) ds) E_{t,T^*}(\exp(-b_2 \int_t^T 1_{\{\tau_A \leqslant s \leqslant \tau_A + \eta\}} ds)) | \mathcal{G}_t)$$

$$= e^{-b_0(T-t)} E(\{e^{-(b_2+\mu)(T-t) - b_1 R_{t,T}} (1 + b_2 \int_t^T e^{-(a_0 - b_2)(x-t) - a_1 R_{t,x}} dx)$$

$$+ \frac{\mu}{\mu + b_2} e^{-b_1 R_{t,T}} + \mu b_2 e^{-b_2 T + a_0 t - b_1 R_{t,T}} \int_0^{T-t} \int_{T-y}^T [e^{(b_2 - a_0)x - a_1 R_{t,x}} dx] e^{-\mu y} dy\} | \mathcal{G}_t). \tag{E.8}$$