Dynamic Portfolio Optimization with Credit Risk

A THESIS PRESENTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY OF IMPERIAL COLLEGE LONDON

AND THE

DIPLOMA OF IMPERIAL COLLEGE

BY

Longjie Jia

DEPARTMENT OF MATHEMATICS

IMPERIAL COLLEGE

180 QUEEN'S GATE, LONDON SW7 2BZ

OCTOBER 2018

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: _____

Copyright

Thesis advisor: Professor Harry Zheng   Longjie Jia

Dynamic Portfolio Optimization with Credit Risk

ABSTRACT

Credit risk, which considers the risk of loss resulting from a counterparty's failure to meet contractual obligations, was not properly studied in the literature of dynamic portfolio optimization problem until the 2008-2009 financial crisis. This thesis is devoted to the dynamic portfolio optimization with credit risk. Two main topics are studied in this thesis.

In the first topic, we consider a utility maximization problem with defaultable stocks and looping contagion risk. We assume that the default intensity of one company depends on the stock prices of itself and other companies, and the default of the company induces immediate drops in the stock prices of the surviving companies. Under such looping contagion risk framework, we prove that the value function is the unique viscosity solution of the HJB equation. We also perform some numerical tests to compare and analyse the statistical distributions of the terminal wealth of log utility and power utility based on two strategies, one using the full information of intensity process and the other a proxy constant intensity process. The numerical tests confirm that modeling looping contagion risk properly is important, especially in financial distressed period.

The second topic is on dynamic portfolio optimization with contingent convertible (CoCo) bond. As a new type of hybrid product, CoCo bond has the interesting feature that converting from debt to equity is contingent. We model the conversion of CoCo bond by reduced-form approach and assume that the conversion intensity is a deterministic function of the coupon rate and the issuing bank's stock price. Theoretically, we construct the viscosity solution representation between the value

Thesis advisor: Professor Harry Zheng   Longjie Jia

function and the corresponding HJB equation. Practically, we compare the performance between investing into CoCo bond and the issuing bank's stock. We analyse the statistical distributions of terminal wealth of log utility and power utility based on these two investment choices. Our simulation results show that, the CoCo bond holders bear more loss than equity holders if conversion occurs. However, investing into CoCo bond gets more profit (mean) while bearing less market risk (volatility) as long as conversion does not occur.

Keywords: dynamic portfolio optimization, looping contagion risk, HJB equation, viscosity solution, contingent convertible bond, statistical comparisons.

To my family.

ACKNOWLEDGMENTS

I would like to extend thanks to the following people, without whom this thesis would not have been possible.

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Harry Zheng, not only for his effort and excellent guidance to my academic research, but also for his patience and always being considerate whenever I feel stressed throughout my PhD study. I am also grateful to him for providing me the opportunity to join the Imperial and CitiGroup program, such that I have the chance to gain valuable industrial experience alongside my research.

It has been such a privilege to have worked in counterparty risk modeling team of Citi Risk Analytics. I have met so many talented and friendly individuals there. Special thanks are due to my direct manager Marco Polenghi, for his kindness and frankness.

I am also indebted to Prof. Damiano Brigo and Dr. Thomas Cass who were the examiners in my Early Stage Assessment and Late Stage Review. I benefited a lot from the discussions with them.

I should also extend my deepest gratitude to my dear friends Xue Lu, Hao Liu, Huang Feng, Luting Li and so on. They make me feel like I have a warm family in London.

Finally, but by no means least, thanks go to my mum and dad for all the unbelievable support. They are the most important people in my life and I hope they are always happy and healthy.

List of figures

LIST OF TABLES

CONTENTS

1

INTRODUCTION

In the financial market, one of the most important problems for an investor is that, endowed with a fixed initial capital and a selection of financial instruments, how to determine the optimal trading strategy over a fixed time horizon.

Development of Portfolio Optimization Theory

The earliest portfolio optimization theory is called modern portfolio theory (MPT) or mean-variance analysis. Economist Harry Markowitz introduced MPT in Markowitz (1952) for which he was later awarded a Nobel Prize in Economics. MPT discusses the discrete-time portfolio optimization problem on how a risk-averse investor constructs a portfolio to minimize the risk of the portfolio given an expected return. The risk of each trading asset is measured by the variance of its return. Mathematically, the objective of the investor is to minimize $w^T \Sigma w$ subject to $\mu^T w = \alpha$ for a given return level α. Here w is a vector of portfolio weights, and Σ, μ are the covariance matrix and expectation of the returns on the assets in the portfolio, respectively.

The problem can be solved by applying the Lagrange Multiplier method. For the simplest one-period MPT model, the investor determines the optimal weights w at the beginning and keeps w fixed. For the multi-periods MPT model, the investor calculates the one-period optimal weights for each single period and re-balance the weights at the beginning of next period, based on the updated estimation of Σ and μ. Although the modern portfolio theory has been subjected to various criticisms due to its simple structure and assumptions, it is still widely used by investment institutions nowadays as the fundamental methodology.

The dynamic (continuous-time) portfolio optimization theory starts from Merton (1969). Robert C. Merton considers an investor who dynamically makes the decision about how much to consume and allocate her wealth between a risky stock and a risk-free bank account, such that she maximizes the expected utility of her terminal wealth. Mathematically, Merton's classical dynamic portfolio optimization problem can be written as

$$\sup_{(\pi,c)} \mathbb{E}\left[\int_0^T e^{-\rho s}U(c_s)ds + e^{-\rho T}U(X_T)\right],$$

where $U(\cdot)$ is the utility function, c_t is the consumption at time t, ρ is a constant discount factor and X_t is the wealth process given by the stochastic differential equation (SDE):

$$dX_t = \left((r + (\mu - r)\pi_t)X_t - c_t\right)dt + \sigma\pi_t X_t dW_t,$$

where r is the risk-free rate, (μ, σ) are the expected return and volatility of the stock and dW_t is the increment of the Brownian motion, i.e. the stochastic term of the SDE. Considering the special case that the utility function is of the constant relative risk aversion (CRRA) form:

$$U(x) = \frac{x^\gamma}{\gamma},$$

where $\gamma > 0$ is a constant representing the investor's risk aversion level, Merton derives the closed-form solution of the optimal control π_t, the optimal consumption c_t

and the objective function (or called value function). The approach used by Merton to get the explicit solution is now called the PDE approach to dynamic programming, which is one of the most important approaches to solve dynamic portfolio optimization problem. Basically, the PDE approach transforms the dynamic portfolio optimization problem into solving a Hamilton-Jacobi-Bellman (HJB) equation by applying the dynamic programming principle (DPP), see Pham (2009).

Since Merton (1969, 1971), research on dynamic portfolio optimization problem has continued to extend and generalize Merton's classical model. One of the main extensions is to include more practical factors like stochastic volatility (Pham (2002), Flemming and Hernandez-Hernandez (2003), Kraft (2005), Noh and Kim (2011), Fouque et al. (2017)), stochastic interest rate (Korn and Kraft (2002), Kraft (2009)) and transaction costs (Davis and Norman (1990), Janecek and Shreve (2004), Liu and Zheng (2016)).

Consequences of Financial Crisis

Before the 2008-2009 financial crisis, most of the models proposed in the literature assume that the risky asset is only exposed to market risk. Only a handful of paper discuss the dynamic portfolio optimization problem with credit risk. Hou and Jin (2002) is the first to take credit risk into the consideration of a dynamic portfolio optimization problem. They derive optimal finite horizon investment strategy for an investor with power utility function, who allocates her wealth among a stock, a defaultable bond and a risk-free bank account. Bielecki and Jang (2006) further investigate the impact of recovery amount to the optimal strategy. The pain of 2008-2009 financial crisis forces people to consider the importance of controlling credit risk in the investment. Since then, many works on dynamic portfolio optimization with credit risk have been enriched into the literature, e.g. Bo et al. (2010), Capponi and Figueroa-Lopez (2011), Callegaro et al. (2012). Similar to Hou and Jin (2002), they investigate the optimal control in a financial market where there are one stock, one defaultable bond and a risk-free bank account. Therefore, the considered credit

risk of the financial market is still a single default event. However, it is clear in the financial crisis that the failure of one company will spread to the entire financial market, thus having direct impacts on the performance of other related companies. From the lesson of financial crisis, it is essential for a portfolio manager to seriously consider the contagion risk. Jiao and Pham (2013) consider a financial market with one stock which jumps downward at the default time of a counterparty. The counterparty is not traded or affected by the trading stock. Bo and Capponi (2016) consider a market consisting of a risk-free bank account, a stock index, and a set of CDSs. The default probability of any underlying firm is determined by the default state of all the firms in the market. Capponi and Frei (2017) introduce an equity-credit portfolio with a market consisting of a risk-free bank account, defaultable stocks, and CDSs referencing these stocks. The default intensities of companies are functions of stock prices and some external factors.

Another significant consequence due to the financial crisis is the stronger regulation applied to financial institutions. For example, the new regulation rule from Basel III increases the common equity tier 1 capital (CET1) ratio of all the banks from 2% (under the old regulation Basel II) to 4.5%. Under such circumstance, the contingent convertible bond (CoCo bond) as a hybrid security, is introduced into the financial world. The mechanism of CoCo bond is different from traditional convertible bond in that the conversion from debt to equity is contingent, which means CoCo bond is automatically converted into equity according to some pre-defined trigger event, e.g. CET1 ratio $< 5\%$. Due to its equity nature, CoCo bond is designed to improve the issuing bank's ability to absorb loss. Since the first CoCo bond issued by Lloyds Banking Group in December 2009, the CoCo bond market size has been increasing rapidly. Up to June 2018, the CoCo market size has reached €180bn.

As the conversion of CoCo bond is contingent, the credit risk an investor faces in the current financial market is not only constrained to default event any more. The conversion risk, which is a new type of credit risk, is attracting more and more atten-

tion. The research of CoCo bond in the current literature is mainly on the pricing. The CoCo bond pricing is addressed by two main modelling approaches – structural approach and reduced-form approach. The structural approach models the capital of the issuing bank and constructs the relation between conversion and the capital process. The work includes Spiegeleer and Schoutens (2012), Corcuera et al. (2013), Brigo, et al. (2015), Leung and Kwok (2015) and Jang, et al. (2018). The reduced-form approach models the conversion by a pure jump model, e.g. Cheridito and Xu (2015), Chung and Kwok (2016). Unfortunately, to the best of our knowledge, there has been no existing research in the literature on dynamic portfolio optimization of CoCo bond.

Organization of the Thesis

The thesis is organized as follows. The first ambition in this thesis is to construct a genuine looping contagion risk model and investigate the optimal trading strategy and corresponding terminal wealth distribution under such framework. Most of the contagion risk work in the current literature models the interaction between defaults of different firms by an exogenous factor, which strongly depends on the historical calibration of factor parameters. From the experience of 2008-2009 financial crisis, the investment based on historical calibration reacts slow to the sudden crush of the financial market.

In Chapter 3, We build a looping contagion risk model, where the market is assumed to have one risk-free savings account, and multiple defaultable stocks. The underlying companies may default and the value of defaulted stock price becomes zero. The default time of any stock is the first jump time of a pure jump process driven by an intensity process that depends on all the surviving stock prices, and the surviving stock prices jump at time of default. Unlike the exogenous factor model, our looping contagion model has the ability to adjust trading strategies automatically based on observed stock prices in the portfolio, thus reacting immediately to the sudden movement of the financial market. We study a terminal wealth utility maximization

problem with general utility functions under this looping contagion framework. The main contributions of this chapter are summarized by the following bullet points:

- For general utility functions, we prove that the value function is the unique viscosity solution of the corresponding HJB equation. To the best of our knowledge, this is the first attempt to study the viscosity solution properties of the value function in the literature of utility maximization with looping contagion risk.

- We perform numerical tests to compare the statistical distributions of terminal wealth of log utility and power utility based on two trading strategies, one uses the full information of intensity process, the other a proxy constant intensity process. The constant intensity process is a special example of exogenous factor model where the constant intensity is estimated from the historical calibration. The numerical examples show that, statistically, the looping contagion risk model and exogenous factor model have similar performance in general market situations. However, in the case of market crush where there are big falls of stock prices at the start of the investment, the terminal wealth based on strategies using looping contagion risk model would have much higher expected return and standard deviation than the one using a constant intensity. Therefore, one may greatly improve the performance of investment if one uses the looping contagion risk model in a financial crisis period.

The second ambition of this thesis is to fill in the gap of CoCo bond investment research in the literature. In Chapter 4, we study a terminal wealth utility maximization problem with general utility functions, where the investor can dynamically allocate her wealth between a risk-free bank account and one CoCo bond. Since CoCo bond is a hybrid credit derivative whose conversion depends on issuing bank's capital ratio, we apply the virtue of looping contagion risk model to the CoCo bond modelling. We model the conversion of CoCo bond by reduced-form approach, where

the conversion intensity is a function of its coupon rate and the issuing bank's stock price. The main contributions of this chapter are summarized by the following bullet points:

- To the best of our knowledge, we are the first in the literature tackling the dynamic portfolio optimization problem of CoCo bond. We extend Duffie and Singleton (1999) approach to derive the closed-form CoCo bond pricing formula with continuous coupon. Then we apply the change of measure technique and derive the dynamic of CoCo bond price under physical measure.

- We prove the pre-conversion value function is the unique viscosity solution of the corresponding HJB equation.

- We compare the performance between investing into CoCo bond and the stock issued by the same bank. Our numerical results show that, the CoCo bond holders bear much more loss than equity holders when conversion occurs. However, investing into CoCo bond gets more profit (mean) while bearing less market risk (volatility) as long as conversion does not occur.

Finally, we conclude this thesis in Chapter 5.

2

Dynamic Portfolio Optimization with Looping Contagion Risk

2.1 Introduction

There has been extensive research in dynamic portfolio optimization and credit risk modelling, both in theory and applications (see Pham (2009), Brigo and Morini (2013), and references therein). Utility maximization with credit risk is one of the important research areas, which is to find the optimal value and optimal control in the presence of possible defaults of underlying securities or names. The early work includes Korn and Kraft (2003) using the firm value structural approach and Hou and Jin (2002) using the reduced form intensity approach. Defaults are caused by exogenous risk factors such as correlated Brownian motions, Ornstein-Uhlenbeck or CIR intensity processes. Bo et al. (2010) consider an infinite horizon portfolio

optimization problem with a log utility and assume that both the default risk premium and the default intensity dependent on an external factor following a diffusion process and show the pre-default value function can be reduced to a solution of a quasilinear parabolic PDE (partial differential equation). Capponi and Figueroa-Lopez (2011) assume a Markov regime switching model and derive the dynamics of the defaultable bond and prove a verification theorem with applications to log and power utilities. Callegaro et al. (2012) consider a wealth allocation problem with several defaultable assets whose dynamics depend on a partially observed external factor process.

Contagion risk or endogenous risk has grown into a major topic of interest as it is clear that the conventional dependence modelling of assets using covariance matrix cannot capture the sudden market co-movements. The failure of one company will have direct impacts on the performance of other related companies. For example, during the global financial crisis of 2008-2009, the default of Lehman Brothers led to sharp falls in stock prices of other investment banks and stock indices such as Dow Jones US Financial Index. Since defaults are rare events, one may have to rely on the market information of other companies or indices to infer the default probability of one specific company. For example, one can often observe in the financial market data that the stock price of one company has negative correlation with the CDS (credit default swap) spread (a proxy of default probability) of another company, see Appendix A . One commonly used contagion risk model is the interacting intensity model (see Jarrow and Yu (2001)) in which the default intensity of one name jumps whenever there are defaults of other names in a portfolio. Contagion risk has great impact on pricing and hedging portfolio credit derivatives (see Gu et al. (2013)).

There is limited research in the literature on dynamic portfolio optimization with contagion risk. Jiao and Pham (2011) consider a financial market with one stock which jumps downward at the default time of a counterparty which is not traded and not affected by the stock and, for power utility, solve the post-default problem

by the convex duality method and show the process defined by the pre-default value function satisfies a BSDE (backward stochastic differential equations). Jiao and Pham (2013) discuss multiple defaults of a portfolio with exponential utility and prove a verification theorem for the value function characterized by a system of BSDEs. Bo and Capponi (2016) consider a market consisting of a risk-free bank account, a stock index, and a set of CDSs. The default of one name may trigger a jump in default intensities of other names in the portfolio, which in turn leads to jumps in the market valuation of CDSs referencing the surviving names and affects the optimal trading strategies. They solve the problem with the DPP (dynamic programming principle) and, for power utility, find the optimal trading strategy on the stock index is Merton's strategy, and those on the CDSs can be determined by a system of recursive ODEs (ordinary differential equations). Capponi and Frei (2017) introduce an equity-credit portfolio with a market consisting of a risk-free bank account, defaultable stocks, and CDSs referencing these stocks. The default intensities of companies are functions of stock prices and some external factors, which provides a genuine looping contagion default structure. For a log utility investor, there exists an explicit optimal strategy which crucially depends on the existence of CDSs in the portfolio, see Remark 2.7.3 for details.

In this chapter we analyse the interaction of market and credit risks and its impact on dynamic portfolio optimization. The market is assumed to have one risk-free savings account, and multiple defaultable stocks in which the underlying companies may default and the value of defaulted stock price becomes zero. The default time of any stock is the first jump time of a pure jump process driven by an intensity process that depends on all the surviving stock prices, and the surviving stock prices jump at time of default. This setup characterizes an investment with multiple stocks that are closely dependent on each other, both endogenously and exogenously. Compared with exogenous factor models in the literature, which strongly depend on the historical calibration of factor parameters, the looping contagion model has the ability to adjust trading strategies automatically based on current stock prices

in the portfolio. We study a terminal wealth utility maximization problem with general utility functions under this looping contagion framework.

The aforementioned papers by Bo and Capponi (2016) and Capponi and Frei (2017) characterize the value function as a solution of the HJB (Hamilton-Jacobi-Bellman) equation and, for power and log utility respectively, find the optimal trading strategies with some implicit unknown functions. For general utilities, it is essentially impossible one may guess a solution form of the HJB equation nor can one apply the verification theorem. In that case, a standard approach to studying the value function is the viscosity solution method. We prove, in addition to the verification theorem, that the value function is the unique viscosity solution of the HJB equation. The result is important as it lays a solid theoretical foundation for numerical schemes to find the value function, in contrast to the verification theorem that requires priori the existence of a classical solution to the HJB equation, which is in general difficult to prove. To the best of our knowledge, this is the first time the viscosity solution properties of the value function are studied and established in the literature of utility maximization with looping contagion risk. This is one of the main contributions of this chapter.

We perform some numerical and robust tests to compare the statistical distributions of terminal wealth of log utility and power utility based on two trading strategies, one uses the full information of intensity process, the other a proxy constant intensity process. These two strategies may be considered respectively the active and passive portfolio investment. The numerical examples show that, statistically, they have similar terminal wealth distributions, but active portfolio investment is more volatile in general. Furthermore, we illustrate the financial insight of the looping contagion model via a similar numerical test, but with different initial stock prices. The numerical test assumes that the constant intensity is estimated from historical calibration window, but there are big falls of stock prices at the start of the investment. The numerical example shows that the terminal wealth based on

strategies using stock dependent intensity would have much higher expected return and standard deviation than the one using a constant intensity. Therefore, one may greatly improve the performance of investment if one uses the information of stock dependent default intensity in a financial crisis period.

The rest of the chapter is organized as follows. In Section 2.2 we introduce the looping contagion model and give two examples to illustrate the model setting. In Section 2.3 we describe the portfolio optimization problem. In Section 2.4, we define the value function and the corresponding HJB equation system. In Section 2.5, we show by a verification theorem that, the solution of HJB equation system coincides with the value function under regularity assumption. In Section 2.6, we show that the value function is the unique viscosity solution of the HJB equation system by Theorems 2.6.2 and 2.6.6. In Section 2.7, we perform numerical and robust tests with statistical distribution analysis for log and power utility. In Section 2.8 we prove the main theorems in this chapter. Section 2.9 concludes the chapter.

2.2   Looping Contagion Model Setting

Let $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$ be a complete probability space satisfying the usual conditions and $(\mathcal{G}_t)_{t \geq 0}$ a filtration to be specified below. Let the market consist of one risk-free bank account with value process $(B_t)_{t \geq 0}$ and interest rate r and N defaultable stocks with price process $(S_t)_{t \geq 0} := (S_t^1, ..., S_t^N)_{t \geq 0}^T$, where a^T is the transpose of a vector a. Let $(\mathcal{F}_t)_{t \geq 0}$ be the filtration generated by N correlated Brownian motions $(W_t)_{t \geq 0} := (W_t^1, ..., W_t^N)_{t \geq 0}^T$, which represents the market information. Let $\tau := (\tau_1, ..., \tau_N)$ be a vector of nonnegative random variables representing the default time of each defaultable stock, defined by

$$\tau_i := \inf \left\{ s \geq t : \int_t^s h_u^i du \geq \mathcal{X}_i \right\},$$

where $(h_t^i)_{t\geq 0}$ is an intensity rate process and $\mathcal{X}_i$ is a standard exponential variable on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and is independent of the filtration $(\mathcal{F}_t)_{t\geq 0}$, which means that τ_i is a totally inaccessible stopping time. We make the further assumption that $\mathcal{X}_i$ is independent of $\mathcal{X}_j$ for $i \neq j$. Under this assumption, the default of each stock is independent.

Let $(\mathcal{H}_t)_{t\geq 0}$ be the filtration generated by the default indicator process $(\mathbb{H}_t)_{t\geq 0} := (H_t^1, ..., H_t^N)_{t\geq 0}^T$ where each of the default process H_t^i is associated with the intensity process $(h_t^i)_{t\geq 0}$ and defined by $H_t^i := \mathbb{I}_{\{\tau_i \leq t\}}$, the indicator function that equals 0 if $\tau_i > t$ and 1 otherwise. Denote the value of indicator process $\mathbb{H}_t$ by z, thus $z \in I := \{0,1\}^N$. The indicator process $\mathbb{H}_t$ can only jump from $z := (z_1, ...z_N)^T$ to its neighbor state $z^i := (z_1, ..., 1 - z_i, ..., z_N)^T$ with rate $(1 - z_i)h_t^i$ for $i \in \{1, ..., N\}$. We denote N_z the number of surviving stocks when $\mathbb{H}_t = z$ and I_z the set of surviving stock numbers.

Finally, let $(\mathcal{G}_t)_{t\geq 0}$ be an enlarged filtration, defined by $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$, which contains both the market information and the default information. The stopping time τ_i defined in above way satisfies the so-called H-hypothesis, which means any $\mathcal{F}$-square integrable martingale is also a $\mathcal{G}$-square integrable martingale (see Bielecki and Rutkowski (2003)), a property we will use later in the proofs. The market model is driven by the following stochastic differential equations (SDEs):

$$\frac{dS_t^i}{S_{t-}^i} = \mu_i dt + \sigma_i dW_t^i - L_i^T d\mathbb{H}_t, \quad \frac{dB_t}{B_t} = rdt,$$

for integer $i \in \{1, ..., N\}$ where μ_i is the growth rates of S^i, respectively, σ_i is the volatility rate. The vector $L_i := (L_{i1}, ..., L_{iN})^T$ represents the default impact of each stock to the ith stock, thus $L_{ii} = 1$.

All coefficients are positive constants to simplify discussions. We assume that the defaults of stocks do not occur at the same time. At default time τ_i the defaultable stock price S^i falls to zero and the other stock price S^j is reduced by a percentage

of L_{ji} for $i \neq j$. We require that $L_{ii} = 1$ and $L_{ij} < 1$ for $i \neq j$. $L_{ji} < 1$ ensures the other stock price S^j does not fall to zero at default time of τ_i. We denote by K a generic constant which may have different values at different places.

Assumption 2.2.1. The intensity process $(h_t^i)_{t \geq 0}$ of the default indicator process $(H_t^i)_{t \geq 0}$ can be represented by $h_t^i = h^i(S_{t-}^z, z)$, a function of surviving stock prices $S_{t-}^z := (S_{t-}^i)_{i \in I_z}$ and the state of default indicator process $\mathbb{H}_{t-} = z$. For simplicity, we denote $h^i(S_{t-}^z, z)$ by $h_z^i(S_{t-})$. We further assume that h_z^i is bounded and continuous in S_{t-}^z for $\forall z \in I$ and $i \in \{1, \ldots, N\}$.

To classify the looping contagion model setting, we give two examples which contain only two stocks in the market, denoted by $(S_t)_{t \geq 0}$ and $(P_t)_{t \geq 0}$.

Example 2.2.1. (*One-sided contagion*) *In this case, $(S_t)_{t \geq 0}$ denotes the price of ETF (exchange-traded-fund) on DJ US Financial Index and $(P_t)_{t \geq 0}$ denotes the price of a US investment bank. We may treat the ETF as default-free and its stock price reflects the whole US banking industry and thus has impact on the performance of the individual bank. Then the model is given by*

$$\frac{dS_t}{S_{t-}} = \mu^S dt + \sigma^S dW_t^S - L^S dH_t, \quad \frac{dP_t}{P_{t-}} = \mu^P dt + \sigma^P dW_t^P - dH_t,$$

where μ^S and μ^P are the growth rates of S and P, respectively, σ^S and σ^P are the volatility rates, and $L^S < 1$ is the percentage loss of the stock S upon the default of stock P. At default time τ the defaultable stock price P falls to zero and the stock price S is reduced by a percentage of L^S. The intensity process $(h_t)_{t \geq 0}$ of the default indicator process $(H_t)_{t \geq 0}$ can be represented by $h_t = h(S_{t-}, P_{t-})$.

Example 2.2.2. (*Looping contagion*) *In this case, both $(S_t)_{t \geq 0}$ and $(P_t)_{t \geq 0}$ denote the prices of single stocks. Then the model is given by*

$$\frac{dS_t}{S_{t-}} = \mu^S dt + \sigma^S dW_t^S - dH_t^S - L^S dH_t^P, \quad \frac{dP_t}{P_{t-}} = \mu^P dt + \sigma^P dW_t^P - L^P dH_t^S - dH_t^P.$$

14

At default time of S (resp. P), the stock price S (resp. P) falls to zero and the stock price P (resp. S) is reduced by a percentage of L^P (resp. L^S). The intensity process $h^S_{(0,0)}(t)$ (resp. $h^P_{(0,0)}(t)$) of the default indicator process $(H^S_t)_{t\geq 0}$ (resp. $(H^P_t)_{t\geq 0}$) can be represented by $h^S_{(0,0)}(t) = h^S_{(0,0)}(S_{t-}, P_{t-})$ (resp. $h^P_{(0,0)}(t) = h^P_{(0,0)}(S_{t-}, P_{t-})$). After the default of S (resp. P), the intensity process $h^P_{(1,0)}(t)$ (resp. $h^S_{(0,1)}(t)$) of the default indicator process $(H^P_t)_{t\geq 0}$ (resp. $(H^S_t)_{t\geq 0}$) can be represented by $h^P_{(1,0)}(t) = h^P_{(1,0)}(P_{t-})$ (resp. $h^S_{(0,1)}(t) = h^S_{(0,1)}(S_{t-})$).

2.3 Description of Portfolio Optimization Problem

An investor dynamically allocates proportions $(\pi^1, \ldots, \pi^N, 1 - \sum_{i=1}^{N} \pi^i)$ of the total wealth into the stocks and the bank account. The admissible control set $\mathcal{A}$ is the set of control processes π that are progressively measurable with respect to the filtration $(\mathcal{G}_t)$ and $\pi_t \in A$ for all $t \in [0, T]$. The set A is defined by

$$A := \left\{ \pi \in O \text{ and } 1 - \sum_{i=1}^{N} L_{ij}\pi^i \geq \epsilon_A \text{ for } \forall j \in \{1, ..., N\} \right\},$$

where O is a bounded set in $\mathbb{R}^N$ and ϵ_A is a positive constant. The dynamics of the wealth process $(X_t)_{t\geq 0}$ is given by

$$\frac{dX_t}{X_{t-}} = \left(r + \pi_t^T D_t \theta \right) dt + \pi_t^T D_t \sigma dW_t - \pi_{t-}^T D_t L d\mathbb{H}_t, \tag{2.1}$$

where

$$D_t := \begin{pmatrix} 1 - H_t^1 & \ldots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \ldots & 1 - H_t^N \end{pmatrix}, \ \theta := \begin{pmatrix} \mu_1 - r \\ \vdots \\ \mu_N - r \end{pmatrix},$$

$$
\sigma := \begin{pmatrix} \sigma_1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & \sigma_N \end{pmatrix}, \ L := \begin{pmatrix} L_{11} & \dots & L_{1N} \\ \vdots & \vdots & \vdots \\ L_{N1} & \dots & L_{NN} \end{pmatrix}.
$$

The matrix-valued process $(D_t)_{t \geq 0}$ is adapted to the filtration $(\mathcal{H}_t)_{t \geq 0}$ and plays the role of removing the defaulted stocks. Even though the admissible control set is still A after default time τ_i, $\pi_t^i = 0$ and is not a variable but a constant. The requirement $1 - \sum_{i=1}^N L_{ij}\pi^i \geq \epsilon_A$ for $\forall j \in \{1, ..., N\}$ ensures that when jth stock defaults, the maximum percentage loss of the wealth does not exceed $1 - \epsilon_A$, in other words, if x is the pre-default wealth, then the post-default wealth is at least $\epsilon_A x$.

Remark 2.3.1. For a given control process $\pi \in \mathcal{A}$, equation (2.1) admits a unique strong solution that satisfies

$$
\sup_{t \in [0,T]} \mathbb{E}\left[X_t^\alpha\right] \leq Kx^\alpha \tag{2.2}
$$

for any $\alpha > 0$. This can be easily verified as $X_t^\alpha = x^\alpha N_t M_t$, where

$$
\begin{aligned}
N_t \ := \ & \exp\left(\alpha \int_0^t \left(r + \pi_u^T D_u \theta\right) du + \frac{1}{2}(\alpha^2 - \alpha) \int_0^t \pi_u^T D_u \Sigma D_u \pi_u du \right. \\
& \left. + \alpha \sum_{j=1}^N \int_0^t \ln\left(1 - \sum_{i=1}^N L_{ij}\pi_{u-}^i\right) dH_u^j\right), \\
M_t \ := \ & \exp\left(\int_0^t \alpha \pi_u^T D_u \sigma dW_u - \frac{1}{2}\alpha^2 \int_0^t \pi_u^T D_u \Sigma D_u \pi_u du\right), \\
\Sigma \ := \ & \begin{pmatrix} (\sigma_1)^2 & \rho_{12}\sigma_1\sigma_2 & \dots & \rho_{1N}\sigma_1\sigma_N \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{N1}\sigma_1\sigma_N & \rho_{N2}\sigma_2\sigma_N & \dots & (\sigma_N)^2 \end{pmatrix}, \\
\pi_u \ := \ & (\pi_u^1, \dots, \pi_u^N)^T.
\end{aligned}
$$

Note that ρ_{ij} is the correlation between Brownian motion W^i and W^j. Since A is a bounded set and $1 - \sum_{i=1}^N L_{ij}\pi^i \geq \epsilon_A$ for $\forall j \in \{1, ..., N\}$, we have $|N_t| < K$, independent of t, and M_t is an exponential martingale, thus $\mathbb{E}[M_t] = 1$, which gives

(2.2).

Our objective is to maximize the expected utility of the terminal wealth, that is,

$$\sup_{\pi \in \mathcal{A}} \mathbb{E}[U(X_T^\pi)],$$

where U is a utility function defined on $[0, \infty)$ and satisfies the following assumption.

Assumption 2.3.2. The utility function U is continuous, non-decreasing, concave, and satisfies $U(0) > -\infty$ and $|U(x)| \leq K(1 + x^\gamma)$ for all $x \in [0, \infty)$, where $K > 0$ and $0 < \gamma < 1$ are constants.

Remark 2.3.3. Power utility $U(x) = (1/\gamma)x^\gamma$, $0 < \gamma < 1$, satisfies Assumption 2.3.2, but not log utility. The value function of log utility can be solved directly due to the wealth process exponential structure, see Capponi and Frei (2017). We solve the log utility optimal controls through HJB equation in Section 2.7, and show that the solution is equivalent to that reported in Capponi and Frei (2017).

2.4 Value Function and Hamilton-Jacobi-Bellman Equation

Depending on the default scenarios, the value function is defined by

$$v_z(t, x, s) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^\pi)|X_t = x, S_t = s, \mathbb{H}_t = z\right]$$

for $(t, x, s) \in [0, T] \times (0, \infty)^{N_z+1}$ and $z \in I$. Note that if h is independent of s, then the value function v_z is a function of t, x only.

Remark 2.4.1. Combining Assumption 2.3.2 and Remark 2.3.1, we have

$$|v_z(t, x, s)| \leq K(1 + x^\gamma).$$

17

For the one-sided contagion model defined in Example 2.2.1, the problem can be naturally split into pre-default case and post-default case. The latter is a standard utility maximization problem as stock P disappears and the post-default value function v_1 is a function of time t and wealth x only, see Pham (2009). We have the following continuity result for the pre-default value function v_0.

Theorem 2.4.2. *For the one-sided contagion model (Example 2.2.1), assume further that h is non-increasing in p, monotone in s and Lipschitz continuous in s, p, and U satisfies $|U(x_1) - U(x_2)| \leq K |x_1 - x_2|^\gamma$ for all $x_1, x_2 \in [0, \infty)$. Then the pre-default value function v_0 is continuous in $(t, x, s, p) \in [0, T] \times [0, \infty) \times (0, \infty)^2$.*

Remark 2.4.3. We assume h is non-increasing in p as intuitively the default probability of one company is non-increasing with its own stock price. We also assume that h is monotone in s as we consider S and P are strongly correlated in the sense that the default probability of stock P is either positively or negatively affected by the stock S. The continuity of pre-default value function for the one-sided contagion model relies on the special structure that there is only one default process in the place. For general looping contagion models, the continuity of the value function is difficult to obtain as the order of multiple jumps is random.

Applying the DPP, one can show that the value function satisfies the following HJB equation:

$$- \sup_{\pi \in A} \mathcal{L}^\pi w_z(t, x, s) = 0 \tag{2.3}$$

for $(t, x, s) \in [0, T) \times (0, \infty)^{N_z+1}$ and $z \in I$ with terminal condition $w_z(T, x, s) = U(x)$, where $\mathcal{L}^\pi$ is the infinitesimal generator of processes S, $\mathbb{H}$ and X with control

π, given by

$$
\begin{aligned}
\mathcal{L}^{\pi} w_z(t, x, s) \;=\; & \frac{\partial w_z}{\partial t} + (r + \theta^T \pi) x \frac{\partial w_z}{\partial x} + \sum_{i \in I_z} \mu_i s_i \frac{\partial w_z}{\partial s_i} \\
& + \frac{1}{2} \pi^T \Sigma \pi x^2 \frac{\partial^2 w_z}{\partial x^2} + \frac{1}{2} \sum_{i \in I_z} \sigma_i^2 s_i^2 \frac{\partial^2 w_z}{\partial s_i^2} \\
& + \sum_{i,j \in I_z, i<j} \rho_{ij} \sigma_i \sigma_j s_i s_j \frac{\partial^2 w_z}{\partial s_i \partial s_j} + \sum_{i \in I_z} \rho_i^T \sigma \pi \sigma_i x s_i \frac{\partial^2 w_z}{\partial x \partial s_i} \\
& + \sum_{i \in I_z} h_z^i(s) \left(w_{z^i} \left(t, x \left(1 - \sum_{j=1}^N L_{ji} \pi^j \right), s^i \right) - w_z \right), \quad (2.4)
\end{aligned}
$$

where $s^i := (s_1(1 - L_{1i}), \ldots, s_j(1 - L_{ji}), \ldots, s_N(1 - L_{Ni}))^T$ for $j \in I_{z^i}$ and $\rho_i := (\rho_{i1}, \ldots, \rho_{ij}, \ldots, \rho_{iN})^T$ for $j \in I_z$. Note that the dimension of s^i is N_{z^i} which is equal to $N_z - 1$ as we have removed the ith defaulted stock.

2.5 VERIFICATION THEOREM

We next give a verification theorem for the value function.

Theorem 2.5.1. *Assume that the function tuple $w := (w_z)_{z \in I}$ where $w_z \in C\big([0,T] \times (0,\infty)^{N_z+1}\big) \cap C^{1,2,\ldots,2}\big([0,T) \times (0,\infty)^{N_z+1}\big)$ for any $z \in I$ solves (2.3) with the terminal condition $w_z(T, x, s) = U(x)$, that w_z satisfies a growth condition $|w_z(t,x,s)| \leq K(1 + x^\gamma)$ for $0 < \gamma < 1$, that the maximum of the Hamiltonian[*] in (2.3) is achieved at $\widehat{\pi}(t, x, s, z)$ in A, and that SDE (2.1) admits a unique strong solution $X_t^{\widehat{\pi}}$ with control $\widehat{\pi}$. Then w_z coincides with the value function v_z and $\widehat{\pi}$ is the optimal control process.*

Proof. For $\forall \pi \in \mathcal{A}$, define a new process

$$
w\left(u, X_u^{t,x,s,z,\pi}, S_u^{t,s}, \mathbb{H}_u\right) := \sum_{\bar{z} \in I} w_{\bar{z}}\left(u, X_u^{t,x,s,z,\pi}, S_u^{t,s}\right) \mathbb{I}_{\{\mathbb{H}_u = \bar{z}\}}
$$

---

[*]The Hamiltonian refers to $\mathcal{L}^{\pi} w_z$.

where $X_u^{t,x,s,z,\pi}$ denotes the wealth process starting with $X_t = x, S_t = s, \mathbb{H}_t = z$ associated with control process π, and $S_u^{t,s}$ denotes the prices of surviving stocks at time u starting with $S_t = s$.

As $w_{\bar{z}}$ is smooth in (t, x, s) for $\forall \bar{z} \in I$, we can apply Ito's formula to w and get for any time $u \in [t, T]$

$$w(u, X_u^{t,x,s,z\pi}, S_u^{t,s}, \mathbb{H}_u) = w_z(t, x, s) + \int_t^u \sum_{\bar{z} \in I} \mathcal{L}^\pi w_{\bar{z}}(\bar{u}, X_{\bar{u}}^\pi, S_{\bar{u}}) \mathbb{I}_{\{\mathbb{H}_{\bar{u}} = \bar{z}\}} \, d\bar{u} + M_u - M_t,$$

where $\mathcal{L}^\pi w_{\bar{z}}$ is defined in (2.4), and M is a local martingale defined by

$$
\begin{aligned}
M_u \quad := \quad & \sum_{\bar{z} \in I} \left(\sum_{i \in I_{\bar{z}}} \int_t^u \sigma_i S_{\bar{u}}^i \frac{\partial w_{\bar{z}}}{\partial s_i} \mathbb{I}_{\{\mathbb{H}_{\bar{u}} = \bar{z}\}} dW_{\bar{u}}^i + \int_t^u \pi_{\bar{u}}^T \sigma X_{\bar{u}}^\pi \frac{\partial w_{\bar{z}}}{\partial x} \mathbb{I}_{\{\mathbb{H}_{\bar{u}} = \bar{z}\}} dW_{\bar{u}} \right) \\
& + \sum_{\bar{z} \in I} \left(\sum_{i \in I_{\bar{z}}} \int_t^u \left(w_{\bar{z}^i} \left(\bar{u}, X_{\bar{u}-} \left(1 - \sum_{j=1}^N L_{ji} \pi_{\bar{u}}^j \right), S_{\bar{u}-}^i \right) - w_{\bar{z}}(\bar{u}, X_{\bar{u}-}, S_{\bar{u}-}) \right) \right. \\
& \left. \left(dH_{\bar{u}}^i - h_{\bar{z}}^i(S_{\bar{u}-}) \mathbb{I}_{\{\mathbb{H}_{\bar{u}} = \bar{z}\}} d\bar{u} \right) \right).
\end{aligned}
$$

Since $w_{\bar{z}}$ satisfies the HJB equation (2.3), we have $\mathcal{L}^\pi w_{\bar{z}} \leq 0$. Define stopping times

$$\widetilde{\tau}_n := \inf \left\{ u \geq t : \left| X_u^{t,x,s,z,\pi} - x \right| + \sum_{i \in I_z} \left| S_u^i - s_i \right| \geq n \right\} \wedge (T - 1/n),$$

then $M_{u \wedge \widetilde{\tau}_n}$ is a martingale due to the boundedness of control set A and values and derivatives of $w_{\bar{z}}$. Letting $u = T$ and taking expectation on both sides, we have

$$\mathbb{E} \left[w \left(\widetilde{\tau}_n, X_{\widetilde{\tau}_n}^\pi, S_{\widetilde{\tau}_n}, \mathbb{H}_{\widetilde{\tau}_n} \right) \right] \leq w_z(t, x, s)$$

with equality if $\pi = \widehat{\pi}$. Next we show that

$$\lim_{n \to \infty} \mathbb{E} \left[w_{\bar{z}} \left(\widetilde{\tau}_n, X_{\widetilde{\tau}_n}^\pi, S_{\widetilde{\tau}_n} \right) \mathbb{I}_{\{\mathbb{H}_{\widetilde{\tau}_n} = \bar{z}\}} \right] = \mathbb{E} \left[w_{\bar{z}} \left(T, X_T^\pi, S_T \right) \mathbb{I}_{\{\mathbb{H}_T = \bar{z}\}} \right] = \mathbb{E} \left[U \left(X_T^\pi \right) \mathbb{I}_{\{\mathbb{H}_T = \bar{z}\}} \right],$$

$$(2.5)$$

for $\forall \bar{z} \in I$. Since $|w_{\bar{z}}(t, x, s)| \leq K(1 + x^\gamma)$, also noting (2.2), we have

$$\mathbb{E}\left[\left|w_{\bar{z}}\left(\widetilde{\tau}_n, X^\pi_{\widetilde{\tau}_n}, S_{\widetilde{\tau}_n}\right) \mathbb{I}_{\{\mathbb{H}_{\widetilde{\tau}_n} = \bar{z}\}}\right|^\alpha\right] \leq K\left(1 + \mathbb{E}\left[\left(X^\pi_{\widetilde{\tau}_n}\right)^{\alpha\gamma}\right]\right) \leq K(1 + x^{\alpha\gamma}) < \infty$$

for any $\alpha > 1$. Since $w_{\bar{z}}\left(\widetilde{\tau}_n, X^\pi_{\widetilde{\tau}_n}, S_{\widetilde{\tau}_n}\right) \mathbb{I}_{\{\mathbb{H}_{\widetilde{\tau}_n} = \bar{z}\}}$ is uniformly integrable, we can take the limit under the expectation to get (2.5). This shows that $\mathbb{E}[U(X^\pi_T)] \leq w_z(t, x, s)$ with equality if $\pi = \widehat{\pi}$. Furthermore, SDE (2.1) admits a unique strong solution by the assumption, therefore, w_z coincides with the value function v_z and $\widehat{\pi}$ is the optimal control process. $\qquad \square$

Remark 2.5.2. For log utility $U(x) = \ln x$, the assumption $U(0) > -\infty$ is not satisfied. However, one may postulate that the value function has a form $w_z(t, x, s) = \ln x + f_z(t, s)$, where f is a solution of a linear PDE, see (2.8). If we assume $f_z \in C\left([0, T] \times (0, \infty)^{N_z + 1}\right) \cap C^{1,2,\ldots,2}\left([0, T) \times (0, \infty)^{N_z + 1}\right)$ and is bounded, then one can show that w_z is indeed the value function with the same proof as that of Theorem 2.5.1 except one change: instead of using $|w_z(t, x, s)| \leq K(1 + x^\gamma)$, which does not hold for log utility, one uses $|w_z(t, x, s)| \leq K(1 + |\ln x|)$. Since

$$\begin{aligned}
\ln X_u &= \ln x + \int_t^u \left(r + \pi_{\bar{u}}^T D_{\bar{u}}\theta - \frac{1}{2}\pi_{\bar{u}}^T D_{\bar{u}}\Sigma D_{\bar{u}}\pi_{\bar{u}}\right)d\bar{u} + \int_t^u \pi_{\bar{u}}^T D_{\bar{u}}\sigma dW_{\bar{u}} \\
&\quad + \sum_{j=1}^N \int_t^u \ln\left(1 - \sum_{i=1}^N L_{ij}\pi_{\bar{u}-}^i\right) dH_{\bar{u}}^j
\end{aligned}$$

for $u \in [t, T]$, we have $\mathbb{E}\left[|\ln X_u|^2\right] \leq K(1 + (\ln x)^2)$, which provides the required uniform integrability property in the proof.

2.6 VISCOSITY SOLUTION REPRESENTATION

The verification theorem assumes the existence of a classical solution of the HJB equation (2.3), which may not be true for the value function v_z. Next we show that the value functions $\{v_z\}_{z \in I}$ is the unique viscosity solution to the PDE system

characterized by (2.3) based on the following definition.

To facilitate discussions of viscosity solution, we define F function by

$$F_z\left(t, x, s, w, \nabla_{(t,x,s)}w_z, \nabla^2_{(x,s)}w_z\right) = -\sup_{\pi \in A} \mathcal{L}^\pi w_z(t, x, s),$$

where $\nabla_{(t,x,s)}w_z \in \mathbb{R}^{N_z+2}$ is the gradient vector of w_z with respect to (t, x, s), and $\nabla^2_{(x,s)}w_z \in \mathbb{R}^{(N_z+1)\times(N_z+1)}$ is the Hessian matrix of w_z with respect to (x, s). w_z and its derivatives are evaluated at (t, x, s). The HJB equation (2.3) is the same as

$$F_z\left(t, x, s, v, \nabla_{(t,x,s)}v_z, \nabla^2_{(x,s)}v_z\right) = 0$$

for $\forall z \in I$.

Definition 2.6.1. (i) $w := (w_z)_{z \in I}$ is a viscosity subsolution of the PDE system (2.3) on $[0, T) \times (0, \infty)^{N+1}$ if

$$F_{\bar{z}}\left(\bar{t}, \bar{x}, \bar{s}, \varphi, \nabla_{(t,x,s)}\varphi_{\bar{z}}, \nabla^2_{(x,s)}\varphi_{\bar{z}}\right) \leq 0$$

for all $\bar{z} \in I$, $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty)^{N_{\bar{z}}+1}$ and test functions $\varphi := (\varphi_z)_{z \in I} \in C^{1,2,\dots,2}\left([0, T) \times (0, \infty)^{N_z+1}\right)$ such that $(w_{\bar{z}})^*(\bar{t}, \bar{x}, \bar{s}) = \varphi_{\bar{z}}(\bar{t}, \bar{x}, \bar{s})$ and $(w_z)^* \leq \varphi_z$ for $\forall z \in I$ on $[0, T) \times (0, \infty)^{N_z+1}$, where $(w_z)^*$ is the upper-semicontinuous envelope of w_z, defined by $(w_z)^*(\bar{t}, \bar{x}, \bar{s}) = \limsup_{(t,x,s)\to(\bar{t},\bar{x},\bar{s})} w_z(t, x, s)$.

(ii) $w := (w_z)_{z \in \{0,1\}^N}$ is a viscosity supersolution of the PDE system (B.1) on $[0, T) \times (0, \infty)^{N+1}$ if

$$F_{\bar{z}}\left(\bar{t}, \bar{x}, \bar{s}, \varphi, \nabla_{(t,x,s)}\varphi_{\bar{z}}, \nabla^2_{(x,s)}\varphi_{\bar{z}}\right) \geq 0$$

for all $\bar{z} \in I$, $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty)^{N_{\bar{z}}+1}$ and test functions $\varphi := (\varphi_z)_{z \in I} \in C^{1,2,\dots,2}\left([0, T) \times (0, \infty)^{N_z+1}\right)$ such that $(w_{\bar{z}})_*(\bar{t}, \bar{x}, \bar{s}) = \varphi_{\bar{z}}(\bar{t}, \bar{x}, \bar{s})$ and $(w_z)_* \geq \varphi_z$ for $\forall z \in I$ on $[0, T) \times (0, \infty)^{N_z+1}$, , where $(w_z)_*$ is the lower-semicontinuous envelope of w_z, defined by $(w_z)_*(\bar{t}, \bar{x}, \bar{s}) = \liminf_{(t,x,s)\to(\bar{t},\bar{x},\bar{s})} w_z(t, x, s)$.

(iii) We say that w is a viscosity solution of the PDE system (2.3) on $[0,T) \times (0,\infty)^{N+1}$ if it is both a viscosity subsolution and supersolution of (2.3).

Based on the above definition, we have the following viscosity solution property for the value function. As the proofs of the following theorems are lengthy, we show these proofs in Section 2.8.

Theorem 2.6.2. *For $z \in I$, the value function v_z is a viscosity solution of the PDE system (2.3) on $[0,T) \times (0,\infty)^{N+1}$, satisfying the growth condition $|v_z(t,x,s)| \leq K(1+x^\gamma)$ for some constant $0 < \gamma < 1$.*

To prove the uniqueness of the viscosity solution, we need to introduce a structure condition on the model.

Assumption 2.6.3. The following inequality holds:

$$J_z(\pi) \leq \frac{K}{\epsilon} \left(|x_1 - x_2|^2 + \sum_{i \in I_z} |s_{1i} - s_{2i}|^2 \right), \quad \forall \pi \in A, z \in I,$$

where

$$
\begin{aligned}
J_z(\pi) \; := \; & \frac{1}{2} \pi^T \Sigma \pi \left(x_1^2 Q_{1,1} - x_2^2 Q'_{1,1} \right) + \frac{1}{2} \sum_{i \in I_z} \sigma_i^2 \left(s_{1i}^2 Q_{k_i,k_i} - s_{2i}^2 Q'_{k_i,k_i} \right) \\
& + \sum_{i,j \in I_z, i<j} \rho_{ij} \sigma_i \sigma_j \left(s_{1i} s_{1j} Q_{k_i,k_j} - s_{2i} s_{2j} Q'_{k_i,k_j} \right) \\
& + \sum_{i \in I_z} \rho_i^T \sigma \pi \sigma_i \left(x_1 s_{1i} Q_{1,k_i} - x_2 s_{2i} Q'_{1,k_i} \right)
\end{aligned}
$$

and matrices Q and Q' satisfy

$$
\begin{pmatrix} Q & 0 \\ 0 & -Q' \end{pmatrix} \leq \frac{3}{\epsilon} \begin{pmatrix} I_{N_z+1} & -I_{N_z+1} \\ -I_{N_z+1} & I_{N_z+1} \end{pmatrix}.
$$

Remark 2.6.4. The dimension of matrices Q and Q' in Theorem 2.6.6 is N_z+1. We use k_i to represent the right index of matrices which corresponds to s_i where $i \in I_z$.

23

The introduction of k_i is to resolve the gap between s index and matrix index. We use a simple example to illustrate the definition of k_i. For example, $I_z := \{3, 4, 6\}$. In this case, there are three surviving stocks in the market, namely s_3, s_4, s_6. The dimension of matrices Q and Q' is 4 (including 3 surviving stocks and the wealth process x). Then $k_3 = 2, k_4 = 3, k_6 = 4$.

Remark 2.6.5. For the simplest case where there are only two defaultable stocks in the market, e.g. Example 2.2.1 and Example 2.2.2, Assumption 2.6.3 holds for $\forall \rho \in (-1, 1)$, as $J_z(\pi)$ can be written as

$$J_z(\pi) = \frac{1}{2} \xi^T \begin{pmatrix} Q & 0 \\ 0 & -Q' \end{pmatrix} \xi + \frac{1}{2}(1 - \rho^2)(\sigma^P)^2 \zeta^T \begin{pmatrix} Q & 0 \\ 0 & -Q' \end{pmatrix} \zeta,$$

where

$$\xi = \left(m^T \pi x_1, \sigma^S s_1, \rho \sigma^P p_1, m^T \pi x_2, \sigma^S s_2, \rho \sigma^P p_2 \right)^T$$

$$\zeta = \left(\pi^P x_1, 0, p_1, \pi^P x_2, 0, p_2 \right)^T$$

and $m = (\sigma^S, \rho \sigma^P)^T$, $n = (\rho \sigma^S, \sigma^P)^T$ and $\pi := (\pi^S, \pi^P)^T$. Using the matrix inequality and simple algebraic calculation, one can show that

$$
\begin{aligned}
J_z(\pi) \leq \ & \frac{3}{2\epsilon} \left((m^T \pi)^2 |x_1 - x_2|^2 + (\sigma^S)^2 |s_1 - s_2|^2 \right. \\
& + \rho^2 (\sigma^P)^2 |p_1 - p_2|^2 \right) + \frac{3}{2\epsilon}(1 - \rho^2)(\sigma^P)^2 \left((\pi^P)^2 |x_1 - x_2|^2 + |p_1 - p_2|^2 \right).
\end{aligned}
$$

By the boundedness of control set A, Assumption 2.6.3 holds for all $\rho \in (-1, 1)$.

The next result states the uniqueness of the viscosity solution.

Theorem 2.6.6. *Let Assumption 2.6.3 hold. Assume the value function v_z, $z \in I_z$, satisfies the terminal condition $v_z(T-, x, s) = U(x)$ and the boundary conditions $(v_z)^*(t, x, s) = (v_z)_*(t, x, s)$ for (x, s) on the boundary of $[0, \infty)^{N_z+1}$. Then v_z is the unique viscosity solution of the PDE system (2.3) on $[0, T) \times (0, \infty)^{N_z+1}$.*

Remark 2.6.7. The condition $(v_z)^*(t, x, s) = (v_z)_*(t, x, s)$ for (x, s) on the boundary is equivalent to the existence of the limit of the value function v_z at boundary points. This condition is needed as the domain of (x, s) variables is $(0, \infty)^{N_z+1}$, not $(-\infty, \infty)^{N_z+1}$, in which case one may impose some polynomial growth conditions on v_z, see Pham (2009), Remark 4.4.8, for further discussions on this point.

Remark 2.6.8. Taking the one-sided contagion model defined in Example 2.2.1 as the simplest case, we show in Appendix B how to construct a modification of the true value function such that Theorem 2.6.6 can be applied.

2.7 NUMERICAL TESTS

In this section, we perform some statistical and robust tests for log and power utilities. We assume that there are two defaultable stocks and one risk-free bank account in the market (Example 2.2.2).

2.7.1 OPTIMAL STRATEGIES FOR LOG UTILITY

For $U(x) = \ln x$, the post-default case $z = (1, 1)$ is investing into the risk-free bank account, thus $\pi^S = \pi^P = 0$ and $v_{(1,1)}(t, x) = \ln x + r(T - t)$. We conjecture that the pre-default value function $v_{(0,0)}(t, x, s, p)$ takes the form

$$v_{(0,0)}(t, x, s, p) = \ln x + f_{(0,0)}(t, s, p), \tag{2.6}$$

and the value function $v_{(1,0)}(t, x, p), v_{(0,1)}(t, x, s)$ respectively take the forms

$$v_{(1,0)}(t, x, p) = \ln x + f_{(1,0)}(t, p), \quad v_{(0,1)}(t, x, s) = \ln x + f_{(0,1)}(t, s). \tag{2.7}$$

Substituting (2.6) and (2.7) into (2.3), we get a linear PDE for $f_{(0,0)}$ depending on the value of $f_{(1,0)}$ and $f_{(0,1)}$:

$$
\begin{aligned}
&\frac{\partial f_{(0,0)}}{\partial t} + b^T(s,p)\mathcal{D}f_{(0,0)} + \frac{1}{2}\operatorname{Tr}\left(\sigma\sigma^T(s,p)\mathcal{D}^2 f_{(0,0)}\right) \\
&- \left(h^S_{(0,0)}(s,p) + h^P_{(0,0)}(s,p)\right) f_{(0,0)}(t,s,p) + r \\
&+ h^S_{(0,0)}(s,p)f_{(1,0)}(t,p(1-L^P)) + h^P_{(0,0)}(s,p)f_{(0,1)}(t,s(1-L^S)) \\
&+ \sup_{\pi \in A} G_{(0,0)}(s,p,\pi) = 0
\end{aligned}
\tag{2.8}
$$

with the terminal condition $f_{(0,0)}(T,s,p) = 0$, where $G_{(0,0)}$ is defined by

$$
\begin{aligned}
G_{(0,0)}(s,p,\pi) \; := \; & -\frac{1}{2}\pi^T\Sigma\pi + \theta^T\pi + h^S_{(0,0)}(s,p)\ln(1-\pi^S - L^P\pi^P) \\
& + h^P_{(0,0)}(s,p)\ln(1-L^S\pi^S - \pi^P),
\end{aligned}
$$

and the other notations are given by

$$
b(s,p) := \begin{pmatrix} \mu^S s \\ \mu^P p \end{pmatrix}, \; \sigma(s,p) := \begin{pmatrix} \sigma^S s & 0 \\ \rho\sigma^P p & \sqrt{1-\rho^2}\sigma^P p \end{pmatrix},
$$

$$
\mathcal{D}f_{(0,0)} := \begin{pmatrix} \frac{\partial f_{(0,0)}}{\partial s} \\ \frac{\partial f_{(0,0)}}{\partial p} \end{pmatrix}, \; \mathcal{D}^2 f_{(0,0)} := \begin{pmatrix} \frac{\partial^2 f_{(0,0)}}{\partial s^2} & \frac{\partial^2 f_{(0,0)}}{\partial s \partial p} \\ \frac{\partial^2 f_{(0,0)}}{\partial s \partial p} & \frac{\partial^2 f_{(0,0)}}{\partial p^2} \end{pmatrix}.
$$

By the same argument, we get a linear PDE for $f_{(1,0)}$:

$$
\begin{aligned}
&\frac{\partial f_{(1,0)}}{\partial t} + \mu^P p\frac{\partial f_{(1,0)}}{\partial p} + \frac{1}{2}(\sigma^P)^2 p^2\frac{\partial^2 f_{(1,0)}}{\partial p^2} - h^P_{(1,0)}(p)f_{(1,0)}(t,p) \\
&+ r + h^P_{(1,0)}(p)r(T-t) + \sup_{\pi \in A} G_{(1,0)}(p,\pi) = 0
\end{aligned}
$$

with the terminal condition $f_{(1,0)}(T,p) = 0$, where $G_{(1,0)}$ is defined by

$$
G_{(1,0)}(p,\pi) := -\frac{1}{2}(\sigma^P)^2(\pi^P)^2 + (\mu^P - r)\pi^P + h^P_{(1,0)}(p)\ln(1-\pi^P).
$$

The PDE associated with $f_{(0,1)}$ can be obtained similarly.

Assume the control constraint set A is given by

$$A := \left\{ \pi \mid a^S \leq \pi^S \leq b^S \text{ and } a^P \leq \pi^P \leq b^P \right\},$$

where $a^S, b^S, a^P, b^P \in \mathbb{R}$ are chosen such that $1 - L^T \pi \geq \epsilon_A$ for $\forall \pi \in A$. We need to solve a constrained optimization problem:

$$\max_{\pi \in A} G_{(0,0)}(s, p, \pi).$$

Since A is compact and $G_{(0,0)}$ is continuous, there exists an optimal solution which satisfies the Kuhn-Tucker optimality condition

$$\begin{cases} \mu^S - r - (\sigma^S)^2 \pi^S - \rho \sigma^S \sigma^P \pi^P - \dfrac{h^S_{(0,0)}(s,p)}{1 - \pi^S - L^P \pi^P} - \dfrac{L^S h^P_{(0,0)}(s,p)}{1 - L^S \pi^S - \pi^P} + \mu_1 - \mu_2 = 0 \\ \mu^P - r - (\sigma^P)^2 \pi^P - \rho \sigma^S \sigma^P \pi^S - \dfrac{L^P h^S_{(0,0)}(s,p)}{1 - \pi^S - L^P \pi^P} - \dfrac{h^P_{(0,0)}(s,p)}{1 - L^S \pi^S - \pi^P} + \mu_3 - \mu_4 = 0 \end{cases} \quad (2.9)$$

and the complementary slackness condition

$$\mu_1(\pi^S - a^S) = 0, \ \mu_2(b^S - \pi^S) = 0, \ \mu_3(\pi^P - a^P) = 0, \ \mu_4(b^P - \pi^P) = 0, \quad (2.10)$$

where $\mu_i \geq 0$, $i = 1, \ldots, 4$, are Lagrange multipliers. Since π^S can only take value either in the interior of interval $[a^S, b^S]$ or one of two endpoints, the same applies to π^P, we have nine possible combinations.

If both π^S and π^P are interior points, then $\mu_i = 0$ for $i = 1, \ldots, 4$ from (2.10). Assuming that there exists a unique solution $\left((\pi^S)^*_{(0,0)}, (\pi^P)^*_{(0,0)} \right)$ of (2.9) such that $(\pi^S)^*_{(0,0)} \in (a^S, b^S)$ and $(\pi^P)^*_{(0,0)} \in (a^P, b^P)$, then $\left((\pi^S)^*_{(0,0)}, (\pi^P)^*_{(0,0)} \right)$ is the optimal control. We can discuss other cases one by one. For example, if $(\pi^S)^*_{(0,0)} = a^S$ and $(\pi^P)^*_{(0,0)} \in (a^P, b^P)$, then $\mu_2 = \mu_3 = \mu_4 = 0$ from (2.10) and $(\pi^P)^*_{(0,0)}$ and μ_1 are solutions of equation (2.9). If solutions do not satisfy $(\pi^P)^*_{(0,0)} \in (a^P, b^P)$ and $\mu_1 \geq 0$, then this case is impossible.

Remark 2.7.1. Applying Kuhn-Tucker optimality condition to $G_{(1,0)}(p, \pi)$, we get

27

the explicit optimal control for $z = (1, 0)$ such that

$$(\pi^S)^*_{(1,0)} = 0, \quad (\pi^P)^*_{(1,0)} = \frac{\mu^P - r + (\sigma^P)^2 - \sqrt{(\mu^P - r - (\sigma^P)^2)^2 + 4(\sigma^P)^2 h^P_{(1,0)}(p)}}{2(\sigma^P)^2},$$

if $(\pi^P)^*_{(1,0)} \in (a^P, b^P)$. Assuming that the Merton optimal control $\pi^P_M := \frac{\mu^P - r}{(\sigma^P)^2} < 1 - \epsilon_A$, then by Taylor expansion we have

$$(\pi^P)^*_{(1,0)} = \pi^P_M - \frac{h^P_{(1,0)}(p)}{(\sigma^P)^2 - (\mu^P - r)} + o\left(h^P_{(1,0)}(p)\right).$$

Thus the optimal control $(\pi^P)^*_{(1,0)}$ is approximately equal to the Merton optimal control with an extra modification term related to the default intensity. The result for $z = (0, 1)$ can be obtained similarly.

Remark 2.7.2. For the one-sided contagion model given in Example 2.2.1, we can apply the same Kuhn-Tucker method and assuming that $c_0 := \sigma^S - \rho \sigma^P L^S \neq 0$, we get the optimal control π^* as

$$(\pi^S)^* = c_1 + c_2(\pi^P)^* \quad \text{and} \quad (\pi^P)^* = \frac{c_3 + c_4 - \text{sgn}(c_0)\sqrt{\Delta(s, p)}}{c_5}, \qquad (2.11)$$

where $\text{sgn}(x)$ is the sign function which equals 1 if $x > 0$ and -1 if $x < 0$, and c_i, $i = 1, \ldots, 5$, are some constants given by

$$c_1 := (\mu^S - r - L^S(\mu^P - r))/(\sigma^S c_0),$$

$$c_2 := (L^S(\sigma^P)^2 - \rho \sigma^S \sigma^P)/(\sigma^S c_0),$$

$$c_3 := \left(\mu^S - r - c_1(\sigma^S)^2\right)(c_2 L^S + 1),$$

$$c_4 := \left(c_2(\sigma^S)^2 + \rho \sigma^S \sigma^P\right)(1 - c_1 L^S),$$

$$c_5 := 2\left(c_2(\sigma^S)^2 + \rho \sigma^S \sigma^P\right)(c_2 L^S + 1),$$

and $\Delta(s, p)$ is given by

$$\Delta(s, p) = (c_3 - c_4)^2 + 2L^S c_5 h(s, p).$$

The explicit optimal control relies on the special structure such that there is only one jump in the model setting. Since the optimal controls $(\pi^S)^*$ and $(\pi^P)^*$ have closed-form expressions (2.11), one can apply Taylor expansion as in Remark 2.7.1 and get

$$(\pi^S)^* = (\pi^S)^*_M - \frac{L^S h(s, p)}{c_S} + o\left(h(s, p)\right), \quad (\pi^P)^* = (\pi^P)^*_M - \frac{h(s, p)}{c_P} + o\left(h(s, p)\right),$$
(2.12)

where $(\pi^S)^*_M$ and $(\pi^P)^*_M$ are the Merton's optimal controls given by

$$(\pi^S)^*_M := \frac{1}{1 - \rho^2} \frac{\mu^S - r}{(\sigma^S)^2} - \frac{\rho}{1 - \rho^2} \frac{\mu^P - r}{\sigma^S \sigma^P}, \quad (\pi^P)^*_M := \frac{1}{1 - \rho^2} \frac{\mu^P - r}{(\sigma^P)^2} - \frac{\rho}{1 - \rho^2} \frac{\mu^S - r}{\sigma^S \sigma^P},$$

and $c_S > 0, c_P > 0$ are constant. Note that we assume $c_S > 0, c_P > 0$ which are necessary conditions for $\left((\pi^S)^*_M, (\pi^P)^*_M\right) \in A$. By (2.12) one can easily find the sensitivity information of the optimal controls to the changes of these parameters in the one-sided contagion model.

Remark 2.7.3. Capponi and Frei (2017) derive explicit optimal trading strategies for log utility investors when there are N stocks and N CDSs for these stocks. Applying Ito's formula to log wealth process and taking expectation, they get

$$\mathbb{E}[\ln X_T] = \ln x + \int_0^T \mathbb{E}[\alpha_t] dt,$$
(2.13)

where $\alpha_t := r + f(\bar{x}) + \sum_{n \in I_z} h_n g_n(y_n)$ and $\bar{x}$ is a vector of dimension N_z such that each component is a linear combination of N_z controls π into stocks and N_z controls ψ into CDSs, and y_n, $n \in I_z$, are similarly defined. To maximize α_t over controls π and ψ, Capponi and Frei (2017) use a clever trick of maximizing $f(\bar{x})$ and $g_n(y_n)$ separately and derive a linear equation system with $2N_z$ equations and $2N_z$ variables

in π and ψ. The explicit optimal controls come from solving the equation system, see equation (B.3) in the E-companion paper of Capponi and Frei (2017).

The success of finding the explicit optimal control in Capponi and Frei (2017) crucially relies on the existence of equal number of CDSs in the model. When there is no CDS in the portfolio as in our case, maximizing $f(\bar{x})$ and $g_n(y_n)$ separately would result in an incompatible system of $2N_z$ equations with N_z variables. It is therefore impossible to get the closed-form optimal control for log utility investors in our looping contagion model by applying Capponi and Frei's technique. In fact, applying Ito's formula to log wealth process and taking expectation in our model, we get

$$\mathbb{E}[\ln X_T] = \ln x + \int_0^T \mathbb{E}[\widetilde{\alpha}_t]dt,$$

where $\widetilde{\alpha}_t := r + \pi_t^T D_t \theta - \frac{1}{2}\pi_t^T D_t \Sigma D_t \pi_t + \sum_{j \in I_z} h_z^j(s) \ln\left(1 - \sum_{i \in I_z} L_{ij}\pi_{t-}^i\right)$. Taking derivatives of $\widetilde{\alpha}_t$ with respect to π would lead to an equation system similar to that in (2.9).

Although the explicit solution of equation system (2.9) is hard to obtain, we may assume by the virtue of Remark 2.7.2 that the optimal control is approximately the sum of Merton's classical optimal control and a correction function:

$$(\pi^S)_{(0,0)}^* = (\pi^S)_M^* + g_S\left(h_{(0,0)}^S(s,p), h_{(0,0)}^P(s,p)\right) + o\left(h_{(0,0)}^S(s,p)\right) + o\left(h_{(0,0)}^P(s,p)\right),$$

$$(\pi^P)_{(0,0)}^* = (\pi^P)_M^* + g_P\left(h_{(0,0)}^S(s,p), h_{(0,0)}^P(s,p)\right) + o\left(h_{(0,0)}^S(s,p)\right) + o\left(h_{(0,0)}^P(s,p)\right),$$

where g_S and g_P are two correction functions. Setting the following benchmark parameter values

$$r = 0.05, \mu^S = 0.1, \mu^P = 0.15, \sigma^S = 0.3, \sigma^P = 0.4, \rho = 0.0, L^S = 0.2, L^P = 0.3$$

and the control set parameters $a^S = -0.75, b^S = 0.75, a^P = -0.75, b^P = 0.75$, we analyze the Merton add-on functions g_S and g_P by numerical study. Note that

$\left((\pi^S)^*_{(0,0)}, (\pi^P)^*_{(0,0)} \right)$ can be obtained by numerically solving equation system (2.9).
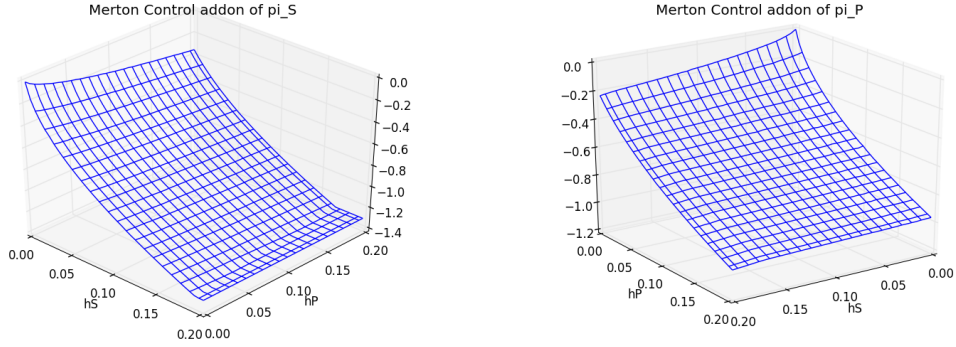
Figure 2.1: Merton add-on functions g_S and g_P w.r.t. intensity values

Figure 2.1 shows the g_S and g_P functions with respect to $h^S_{(0,0)}(s,p)$ and $h^P_{(0,0)}(s,p)$. The Merton optimal controls in above case are $(0.56, 0.62)$. The first observation from Figure 2.1 is that the add-on functions g_S and g_P are always non-positive. It makes sense as intuitively investors hold less (or short sell more) of the risky assets when there is default risk in the place. The second observation is that the absolute value of g_S (resp. g_P) increases with both $h^S_{(0,0)}(s,p)$ and $h^P_{(0,0)}(s,p)$. However, it seems that g_S (resp. g_P) is more sensitive to the change of $h^S_{(0,0)}(s,p)$ (resp. $h^P_{(0,0)}(s,p)$).

2.7.2 Performance comparison of state-dependent and constant intensities

We now do some statistical analysis. The data used are the same as the benchmark case and:

$$T = 1, S_0 = 100, P_0 = 100, x_0 = 100.$$

Assume the intensity function h is given by

$$h(x,y) = \min\left\{\max\left\{h_0\left(k_1 x + k_2 y\right)^{-\alpha}, h_m\right\}, h_M\right\} \qquad (2.14)$$

with minimum intensity $h_m = 0.05$, maximum intensity $h_M = 1.0$, and parameter $\alpha = 1$. The default intensity functions with respect to each stock and default state are given by

$$h^S_{(0,0)}(s,p) = h(s,p), \quad h^P_{(0,0)}(s,p) = h(p,s), \quad h^S_{(0,1)}(s) = h(s,0), \quad h^P_{(1,0)}(p) = h(p,0).$$

Note that h_0 controls the initial intensity and weights k_1, k_2 control the sensitivity of intensity h to stock prices s and p. We set $h_0 = 10.0$ such that the initial intensity is 0.1 and $k_1 = 0.7, k_2 = 0.3$ which means the default intensity of one stock is slightly more sensitive to its own stock price. Moreover, the intensity of one stock jumps up when the other stock defaults, which captures the virtue of interacting default intensity model, see Bo and Capponi (2013).
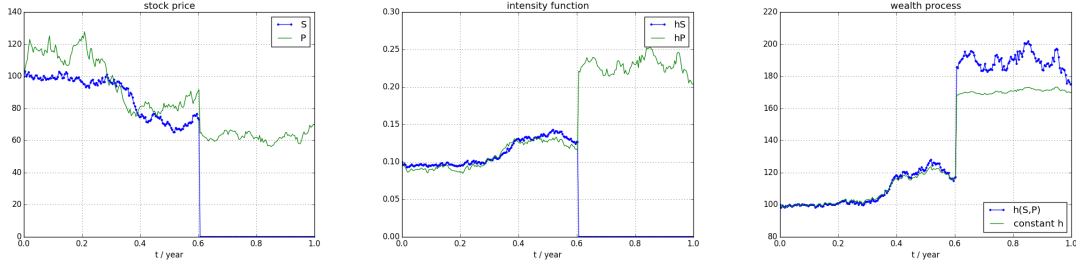
Figure 2.2: Sample paths of stock price, default intensity, and wealth (looping contagion)

Figure 2.2 shows sample paths of stock prices, default intensities, and optimal wealth with two different trading strategies. The left panel shows stock price paths of S and P. In this scenario, only stock S defaults. At time of default, stock price S drops to zero and stock price P jumps down then continues. The middle panel shows the default intensity processes $h^S_z(S_t, P_t)$ and $h^P_z(S_t, P_t)$, which are functions of stock prices S_t, P_t. The intensity of stock S becomes zero after default, while the default intensity $h^P_{(0,0)}(S_t, P_t)$ jumps up to $h^P_{(1,0)}(P_t)$. The right panel shows the sample wealth paths when optimal control strategies used are based on S, P-dependent

intensities and constant intensities (value equal to 0.1). Both the wealth paths with S, P-dependent intensities and constant intensities jump up when default occurs and then two wealth paths move in the same pattern. Compared with the constant intensities, the S, P-dependent wealth path jumps more. This is not surprising as at time of default, strategies with intensity $h(S_t, P_t)$ short sells more stocks S and P than strategies with constant intensity, which means gain is more, see Figure 2.3. Of course, this is due to the fact that at time of default the default intensities of S and P are both above 0.1. The opposite phenomenon happens when the intensity $h^P(S_t, P_t)$ at time of default is larger than constant intensity 0.1.
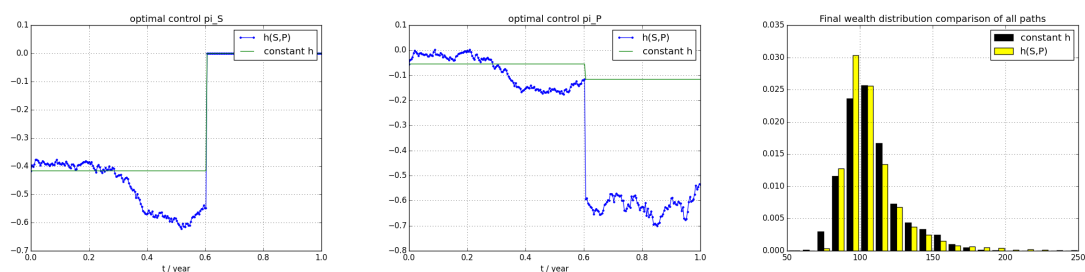
Figure 2.3: Optimal controls and terminal wealth distribution (looping contagion)

Figure 2.3 shows optimal controls π^S and π^P associated with the stock paths in Figure 2.2 and the statistical distributions of the wealth at time T. The left panel and mid panel are proportions of wealth invested in stocks S and P, respectively. It is clear that as default intensity increases, investments in stocks S and P both decrease and investment in savings account B increases, which is intuitively expected as if the default probability of one stock increases, then one would reduce the holdings of both stocks S and P to reduce the risk of loss in case the default of system indeed occurs. In this scenario, both the optimal investment strategies to stock S and P are short-selling, which is a combination effect of parameters chosen and default in the place. We simulate 10000 paths of both stock prices S and P, using S, P-dependent default intensity $h(S_t, P_t)$. Among all these paths, about $1/5$ (precisely 1752 paths) contain defaults of either S or P. The terminal wealth is generated by two strategies: one is optimal strategy based on the full information

of $h(S_t, P_t)$, the other is optimal strategy based on constant intensity 0.1. The right panel shows the histograms of terminal wealth of these two strategies. It is clear that their distributions are similar but with some slight differences at tail parts, that is, probability of over-performance is higher and probability of under-performance is lower with S, P-dependent intensities. These histograms seem to indicate, for log utility, the overall performance of S, P-dependent optimal strategies and constant strategies are similar, while the S, P-dependent optimal strategies perform better in extreme scenarios.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + $h(S, P)$	107.78	22.60	84.08	171.82
All samples + constant h	107.59	19.27	78.60	157.80
Default + $h(S, P)$	134.81	36.04	83.80	225.20
Default + constant h	134.68	21.50	95.29	178.04
No-default + $h(S, P)$	102.17	12.82	84.11	134.68
No-default + constant h	101.96	12.99	78.13	130.48

Table 2.1: Sample means, standard deviations, and quantile values (looping contagion)

Table 2.1 contains sample means, sample standard deviations, and quantile values at low end (2.3%) and high end (97.7%) for both S, P-dependent intensity and constant intensity 0.1. It is clear that the overall sample mean with S, P-dependent optimal strategies is (slightly) higher than constant intensity optimal strategies, which is expected as the former is the genuine optimal control, however, the sample standard deviation with S, P-dependent optimal strategies is also higher, which implies the S, P-dependent optimal strategies can be volatile and risky, while the constant optimal strategies are more conservative. However, if we check the quantiles of the distribution (which is a different risk measure), we find that the S, P-dependent optimal strategies overall generate both higher 2.3% quantile (less loss) and higher 97.7% quantile (more gain), which implies the S, P-dependent optimal strategies outperform the constant strategies in the extreme scenarios. Note that the outperformance in upper quantile comes from more short selling (anticipating the default when stock price is very low).

Remark 2.7.4. By far the conclusion drawn relies on the benchmark parameter values, in which case the optimal controls for both stocks are short selling in most scenarios. We repeat the same comparison tests on two other parameter sets (if not specified, the parameter value is the same as benchmark case).

- Parameter set 1. $\sigma^S = 0.2, \sigma^P = 0.3, L^S = 0.1, L^P = 0.2, \rho = 0.4, h_0 = 5.0, h_m = 0.01$.

- Parameter set 2. $r = 0.01, \mu^S = 0.15, \mu^P = 0.2, L^S = 0.05, L^P = 0.1, \rho = 0.7, h_0 = 5.0, h_m = 0.01$.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + $h(S,P)$	106.79	12.52	73.02	133.62
All samples + constant h	106.71	11.21	75.90	128.36
Default + $h(S,P)$	96.66	23.41	67.86	136.54
Default + constant h	99.04	23.89	68.48	136.76
No-default + $h(S,P)$	107.78	10.37	92.16	132.88
No-default + constant h	107.46	8.72	91.44	125.64

Table 2.2: Sample means, standard deviations, and quantile values of Parameter set 1

Table 2.2 contains the statistics of parameter set 1 for both S, P-dependent intensity and constant intensity 0.05. In most scenarios, the strategies of parameter set 1 is short selling stock S and longing stock P. As the initial default intensity in this case is 0.05, there are 887 paths containing default among the 10000 simulated paths.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + $h(S,P)$	111.27	33.75	56.86	195.32
All samples + constant h	110.65	29.87	54.83	177.62
Default + $h(S,P)$	64.88	12.10	50.69	97.75
Default + constant h	66.44	16.80	40.16	103.75
No-default + $h(S,P)$	115.72	31.75	72.29	198.58
No-default + constant h	114.90	27.30	69.26	179.72

Table 2.3: Sample means, standard deviations, and quantile values of Parameter set 2

Table 2.3 contains the statistics of parameter set 2 for both S, P-dependent intensity and constant intensity 0.05. In most scenarios, the strategies of parameter set 2 are longing both stocks S, P. There are 876 paths containing default among the 10000 simulated paths.

From Table 2.1, 2.2, 2.3, the overall performance with S, P-dependent optimal strategies is very similar to that with constant intensity optimal strategies. The overall sample mean with S, P-dependent optimal strategies is (slightly) higher than constant intensity optimal strategies, and the sample standard deviation with S, P-dependent optimal strategies is also higher, which implies the difference in the tail distribution and that S, P-dependent optimal strategies can be more volatile.

2.7.3 ROBUST TESTS OF MODEL PARAMETERS

Assume intensity function h is given by (2.14) and stock prices S and P are generated based on that. It may be difficult to calibrate parameters accurately even one knows the exact form of the intensity function. We do some robust tests for parameters $k_1, k_2, \alpha, h_m, h_M, h_0$, that is, we compare the optimal performances of two investors, one uses benchmark parameter values and the other incorrect estimated values. We change one parameter only in each test while keep all other parameters fixed at benchmark values.

	mean	std dev	2.3% quantile	97.7% quantile
BM	107.78	22.60	84.08	171.82
k case1	107.69 (-0.09%)	22.85 (1.10%)	83.00 (-1.29%)	173.82 (1.16%)
k case2	107.69 (-0.09%)	23.15 (2.44%)	80.95 (-3.73%)	174.67(1.66%)
$\alpha = 0.8$	111.59 (3.53%)	53.15 (135.20%)	53.18 (-36.75%)	262.58 (52.82%)
$\alpha = 1.2$	105.17 (-2.42%)	9.43 (-58.26%)	84.52 (0.53%)	122.93 (-28.45%)
h case1	107.77 (-0.01%)	22.60 (0.01%)	84.08 (0.00%)	171.82 (0.00%)
h case2	107.76 (-0.02%)	22.62 (0.07%)	83.91 (-0.20%)	171.82 (0.00%)
$h_0 = 5$	105.36 (-2.25%)	9.52 (-57.86%)	85.24 (1.39%)	124.36 (-27.62%)
$h_0 = 15$	109.76 (1.84%)	37.64 (66.55%)	70.59 (-16.04%)	221.49 (28.91%)

Table 2.4: Robust test of intensity parameters (looping contagion)

Note that BM represents the benchmark case. For the other notations, k case1: $k_1 = 0.5, k_2 = 0.5$; k case2: $k_1 = 0.3, k_2 = 0.7$; h case1: $h_m = 0.01, h_M = 1.5$; h case2: $h_m = 0.07, h_M = 0.5$.

Table 2.4 shows that sample means are essentially the same over a broad range of model parameters. The main difference is sample standard deviations. Percentage changes over the benchmark values are listed in parentheses. The performance of state-dependent intensity strategies is robust for some parameters, including weight k_1, k_2, minimum intensity level h_m and maximum intensity level h_M. Changes of these parameters do not greatly change sample standard deviations and quantile values at low and high ends. On the other hand, it seems important to have correct estimations of parameters α and h_0 to avoid large changes of the standard deviation. Those parameters have strong impact on the estimated intensity levels. For example, if one overestimates the initial default intensity ($h_0 = 15$ instead of correct value $h_0 = 10$) then the sample standard deviation is greatly increased with large loss at low end quantile value.

Next we do some robust tests to see the impact of changes of model parameters on the distribution of optimal terminal wealth, including drift μ, volatility σ, correlation ρ, and percentage loss L^S. We change drift and volatility parameters by 20% of their benchmark values and correlation and percentage loss parameters by some big deviations.

Table 2.5 lists statistical results of distributional sensitivity to changes of parameters. It is clear that sample means are essentially the same for all parameters, but sample standard deviations are sensitive to changes of drift, volatility, correlation and percentage loss, which would significantly affect overall distributions of optimal terminal wealth. This requires one to have good estimations of these parameters to have correct distributions. It is well known that it is easy to estimate volatility but difficult to estimate drift (see Rogers (2013)) and information of percentage loss is rarely available. Since optimal trading strategies and optimal wealth distributions

37

	mean	std dev	2.3% quantile	97.7% quantile
benchmark	107.78	22.60	84.08	171.82
$\mu^S = 0.12$	107.05 (-0.68%)	17.54 (-22.39%)	91.98 (9.39%)	158.13 (-7.97%)
$\mu^S = 0.08$	108.55 (0.71%)	28.35 (25.44%)	75.99 (-9.62%)	187.88 (9.35%)
$\mu^P = 0.18$	107.53 (-0.23%)	21.78 (-3.61%)	80.93 (-3.75%)	165.05 (-3.94%)
$\mu^P = 0.12$	108.08 (0.28%)	25.58 (13.17%)	83.68 (-0.48%)	182.54 (6.24%)
$\sigma^S = 0.36$	107.28 (-0.46%)	18.87 (-16.49%)	88.04 (4.71%)	161.89 (-5.78%)
$\sigma^S = 0.24$	108.47 (0.64%)	27.88 (23.37%)	78.37 (-6.79%)	188.35 (9.62%)
$\sigma^P = 0.48$	107.74 (-0.04%)	21.94 (-2.92%)	84.12 (0.04%)	169.01 (-1.63%)
$\sigma^P = 0.32$	107.86 (0.07%)	23.56 (4.24%)	83.84 (-0.29%)	175.82 (2.33%)
$\rho = -0.3$	108.21 (0.40%)	25.50 (12.83%)	83.45 (-0.75%)	183.56 (6.83%)
$\rho = 0.3$	107.57 (-0.20%)	21.83 (-3.39%)	82.74 (-1.59%)	167.26 (-2.65%)
$L^S = 0.1$	107.49 (-0.26%)	20.49 (-9.33%)	87.32 (3.85%)	166.49 (-3.10%)
$L^S = 0.4$	108.24 (0.43%)	26.45 (17.05%)	77.52 (-7.80%)	181.95 (5.90%)
$L^P = 0.15$	107.70 (-0.07%)	21.94 (-2.91%)	83.26 (-0.97%)	169.43 (-1.39%)
$L^P = 0.6$	107.88 (-0.10%)	23.90 (5.74%)	84.69 (0.73%)	177.53 (3.32%)

Table 2.5: Robust test of model parameters (looping contagion)

are greatly influenced by these parameters which are difficult to be correctly esti-
mated, one needs to be cautious in using state-dependent intensity to model and
solve optimal investment problems. Using sub-optimal but conservative and robust
trading strategies, instead of optimal ones based on unobservable parameters and
intensities, might be more sensible and less risky.

2.7.4 PERFORMANCE COMPARISON OF DIFFERENT INITIAL STOCK PRICES

Table 2.1 shows the overall distributions of the terminal wealth are similar whether
one uses the intensity $h_z^i(s, p)$ or constant intensity 0.1 as approximation. This is
possibly due to the fact that the initial price of S and P are both 100, which results
in the inital intensity $h_z^i(s, p)$ being equal to the constant intensity. The value 0.1
comes from the calibration which relies only on the historical data, while $h(s, p)$
is a forward-looking function which depends on the future stock prices. Table 2.1

represents the normal situation where the default probability in calibration window is close to that in investment window. However, if the initial intensity $h_z^i(s,p)$ is vastly different from 0.1 which comes from the estimation of calibration window (one example is that the calibration window is just before the financial crisis, while the investment starts from the financial crisis period), the distributions of terminal wealth can be significantly different. We use a numerical example to illustrate this.

For simplicity, let the intensity function be given by $h_z^i(s,p) = 20/(s+p)$. This means the default intensity of S jumps from $20/(S_t + P_t)$ to $20/S_t$ after P defaults. So is the situation when S defaults. Assume that the initial prices of S and P are $s = 10, p = 10$ respectively, then the initial intensity is $h_{(0,0)}^S(s,p) = h_{(0,0)}^P(s,p) = 1$, which makes the stocks ten times more likely to default than the constant intensity $h = 0.1$ would have suggested (from the calibration window). This would cause one to take different control strategies (more shortselling when $s = 10, p = 10$) and would have large impact on the distributions of the terminal wealth as shown in the table below.

	mean	std dev	2.3% quantile	97.7% quantile
All samples & $h(S,P)$	179.74	63.21	58.63	316.08
All samples & $h \equiv 0.1$	66.26	22.87	44.62	138.28
Default & $h(S,P)$	195.90	52.78	99.20	318.65
Default & $h \equiv 0.1$	57.70	7.16	44.23	73.82
No-default & $h(S,P)$	88.18	31.46	50.36	172.72
No-default & $h \equiv 0.1$	114.78	20.69	78.99	157.11

Table 2.6: Sample statistics with different initial stock prices. Data: $S_0 = 10$, $P_0 = 10$

Table 2.6 shows the statistics of the terminal wealth with 10000 simulation scenarios which produces 8542 default scenarios, a reflection of the high initial default intensity $h_z^i(S_0, P_0) = 1$. When stock prices are small, defaultable stocks are very likely to default. With S, P dependent intensity, the optimal controls are to short sell more stocks, which results in a much larger mean (195.90) than the mean (57.70) with constant intensity $h \equiv 0.1$ if stock S or P indeed defaults (anticipated). However, if

stock does not default (non-anticipated), then the opposite outcomes appear. This numerical test shows S, P-dependent control strategies may outperform or underperform S, P independent control strategies, depending on the anticipated market event (default of stock) occurring or not.

Remark 2.7.5. We repeat the same tests on the other two parameter sets defined in Remark 2.7.4. Table 2.7 and Table 2.8 show the statistics of the other two parameter sets respectively. It is clear that the conclusion drawn in this section is very robust.

	mean	std dev	2.3% quantile	97.7% quantile
All samples & $h(S, P)$	165.68	52.40	65.29	273.71
All samples & $h \equiv 0.1$	52.50	32.25	29.15	151.40
Default & $h(S, P)$	179.67	43.00	100.40	277.57
Default & $h \equiv 0.1$	40.00	7.11	28.53	57.25
No-default & $h(S, P)$	88.33	25.18	57.77	158.67
No-default & $h \equiv 0.1$	121.61	29.54	75.94	191.41

Table 2.7: Sample statistics with different initial stock prices of Parameter set 1

	mean	std dev	2.3% quantile	97.7% quantile
All samples & $h(S, P)$	141.65	57.07	54.22	260.60
All samples & $h \equiv 0.1$	53.79	42.91	24.60	197.84
Default & $h(S, P)$	154.05	52.26	65.22	268.57
Default & $h \equiv 0.1$	38.37	9.68	24.39	62.05
No-default & $h(S, P)$	73.10	24.79	52.71	143.56
No-default & $h \equiv 0.1$	139.10	53.96	65.75	265.55

Table 2.8: Sample statistics with different initial stock prices of Parameter set 2

2.7.5 NUMERICAL METHOD FOR POWER UTILITY

For power utility $U(x) = (1/\gamma)x^\gamma$, $0 < \gamma < 1$, the post-default case is well known with the optimal control $\pi^S = (\mu^S - r)/((\sigma^S)^2(1 - \gamma))$ (and $\pi^P = 0$) and the post-

default value function $v_1(t, x) = (1/\gamma)x^\gamma g_1(t)$, where

$$g_1(t) = \exp\left(\left(r\gamma + \frac{\gamma}{2(1-\gamma)}\left(\frac{\mu^S - r}{\sigma^S}\right)^2\right)(T-t)\right).$$

We conjecture that the pre-default value function takes the form

$$w(t, x, s, p) = \frac{x^\gamma}{\gamma}f(t, s, p). \tag{2.15}$$

Substituting (2.15) into (2.3), we get a semilinear PDE for f:

$$-\frac{\partial f}{\partial t} - \frac{1}{2}\text{Tr}\left(\sigma\sigma^T(s, p)\mathcal{D}^2 f\right) - \sup_{\pi \in A}\left\{b^T(s, p, \pi)\mathcal{D}f - \beta(s, p, \pi)f + g(t, s, p, \pi)\right\} = 0, \tag{2.16}$$

with terminal condition $f(T, s, p) = 1$, where

$$b(s, p, \pi) := \begin{pmatrix} \left(\mu^S + \gamma m^T \pi \sigma^S\right)s \\ \left(\mu^P + \gamma n^T \pi \sigma^P\right)p \end{pmatrix}, \quad \sigma(s, p) := \begin{pmatrix} \sigma^S s & 0 \\ \rho\sigma^P p & \sqrt{1-\rho^2}\sigma^P p \end{pmatrix},$$

$$\mathcal{D}f := \begin{pmatrix} f_s \\ f_p \end{pmatrix}, \quad \mathcal{D}^2 f := \begin{pmatrix} f_{ss} & f_{sp} \\ f_{sp} & f_{pp} \end{pmatrix},$$

and

$$\beta(s, p, \pi) := -r\gamma + h(s, p) - \gamma\left(\theta^T \pi + \frac{1}{2}(\gamma - 1)\pi^T \Sigma \pi\right),$$

$$g(t, s, p, \pi) := h(s, p)g_1(t)(1 - L^T\pi)^\gamma.$$

Equation (2.16) is a nonlinear PDE with two state variables and it is highly unlikely, if not impossible, to find a closed form solution f. However, by Pham (2009) (Remark 3.4.2), equation (2.16) is the HJB equation for the value function v of the following optimal control problem:

$$v(t, y) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[\int_t^T \Gamma(t, u)g(u, Y_u, \pi_u)du + \Gamma(t, T)\bigg| Y_t = y\right], \tag{2.17}$$

41

where $Y_u := (S_u, P_u)^T$, $t \leq u \leq T$, is a controlled Markov state process satisfying the following SDE:

$$dY_u = b(Y_u, \pi_u)du + \sigma(Y_u)dW_u, \ t \leq u \leq T, \tag{2.18}$$

with the initial condition $Y_t = y := (s, p)^T$, W is a 2-dimensional standard Brownian motion and $\Gamma(t, u) := \exp\left\{-\int_t^u \beta(Y_l, \pi_l)dl\right\}$ is a discount factor.

By our theoretical result, we claim that the value function $v(t, y)$ is the unique viscosity solution of the HJB equation (2.16). Moreover, if the HJB equation (2.16) has a classical solution, then it is the value function $v(t, y)$. In other words, we may find the solution $f(t, s, p)$ of equation (2.16) by solving a stochastic optimal control problem (2.17). Since the diffusion coefficient of SDE (2.18) does not contain control variable π, we may use the numerical method of Kushner and Dupuis (2001) to find the optimal value function in (2.17), which would give us a numerical approximation to the solution $f(t, s, p)$ of equation (2.16). Next we give some details.

According to Kushner and Dupuis (2001), the process Y can be approximated by a Markov chain process, which transits a point $Y_t = (s, p)$ at time t to one of nine points $Y_{t+\Delta t}$ may take at time $t + \Delta t$, that is, $(s, p), (s \pm \delta, p), (s, p \pm \delta), (s + \delta, p \pm \delta), (s - \delta, p \pm \delta)$, with the following transition probabilities:

$$
\begin{aligned}
a^{\delta, \Delta t}\left((s, p), (s, p) \mid \pi\right) &:= 1 - \frac{\Delta t}{\delta}(|b_1| + |b_2|) \\
&\quad - \frac{\Delta t}{\delta^2}\left((\sigma^S s)^2 + (\sigma^P p)^2 - |\rho|\sigma^S \sigma^P sp\right) \\
a^{\delta, \Delta t}\left((s, p), (s \pm \delta, p) \mid \pi\right) &:= \frac{\Delta t}{\delta}b_1^{\pm} + \frac{\Delta t}{2\delta^2}(\sigma^S s)^2 - \frac{\Delta t}{2\delta^2}|\rho|\sigma^S \sigma^P sp \\
a^{\delta, \Delta t}\left((s, p), (s, p \pm \delta) \mid \pi\right) &:= \frac{\Delta t}{\delta}b_2^{\pm} + \frac{\Delta t}{2\delta^2}(\sigma^P p)^2 - \frac{\Delta t}{2\delta^2}|\rho|\sigma^S \sigma^P sp \\
a^{\delta, \Delta t}\left((s, p), (s + \delta, p \pm \delta) \mid \pi\right) &:= \frac{\Delta t}{2\delta^2}\rho^{\pm}\sigma^S \sigma^P sp \\
a^{\delta, \Delta t}\left((s, p), (s - \delta, p \mp \delta) \mid \pi\right) &:= \frac{\Delta t}{2\delta^2}\rho^{\pm}\sigma^S \sigma^P sp,
\end{aligned}
$$

where δ is the step size of space, $\Delta t := (T-t)/N$ is the step size of time with $N \geq 1$ an integer, $b_1 := (\mu^S + \gamma m^T \pi \sigma^S)s$, $b_2 := (\mu^P + \gamma n^T \pi \sigma^P)p$ and $x^+ := \max\{x, 0\}$, $x^- := \max\{-x, 0\}$.

The numerical scheme is based on the following discretized dynamic programming principle:

$$
\begin{aligned}
v(k\Delta t, & S_{k\Delta t}, P_{k\Delta t}) \\
\approx \quad & \sup_{\pi_k \in A} \bigg(g(k\Delta t, S_{k\Delta t}, P_{k\Delta t}, \pi_k)\Delta t \\
& + \exp\left\{-\beta(S_k^N, P_k^N, \pi_k)\Delta t\right\} \mathbb{E}\left[v\left((k+1)\Delta t, S_{k+1}^N, P_{k+1}^N\right)\right] \bigg)
\end{aligned}
$$

for $k = N-1, \ldots, 1, 0$, where π_k is the piece-wise constant control and the expectaton is computed with the help of the above Markov chain transition probabilities. The terminal condition is given by $v(N\Delta t, S_{N\Delta t}, P_{N\Delta t}) = 1$.

We compare the passive investment and active investment under the power utility setting. Most parameter values used in power utility case are the same as log utility benchmark case, except the step size of space $\delta = 5$, the step size of time $\Delta t = 0.1$, and the set of control parameters are $a^S = a^P = -1.0, b^S = b^P = 1.0$. Table 2.9 lists the numerical results with mean, variance, and quantile values at lower and upper ends. It is clear the performance is similar to that of the log utility as one would expect.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + $h(S, P)$	106.38	17.45	77.30	148.20
All samples + constant h	105.99	14.90	79.92	139.08
Default + $h(S, P)$	106.70	21.00	71.67	158.63
Default + constant h	106.18	15.78	78.50	140.40
No-default + $h(S, P)$	106.35	17.06	77.83	146.52
No-default + constant h	105.97	14.80	80.03	138.94

Table 2.9: Sample means, standard deviations, and quantile values with power utility (looping contagion)

There is a backward stochastic differential equation (BSDE) representation of the solution $f(t, s, p)$ of equation (2.16). So in theory one may find f if one can solve a highly nonlinear BSDE, which is not pursued in this thesis, see Cheridito et al. (2007) for details.

Remark 2.7.6. Solving the optimal control for power utility relies on the numerical method in Kushner and Dupuis (2001). By guessing the specific form of the value function (2.15), we simplify the HJB equation (2.16) such that the resulting controlled process (2.18) has volatility term $\sigma(Y_u)$ independent of control π. For general utility function, the controlled process is normally dependent of π, thus the convergence result of Kushner and Dupuis (2001) is invalid. For the completeness of the numerical test section, we develop a numerical method in Appendix C, which is adapted to the general utility function.

2.8   PROOFS OF THE MAIN THEOREMS

We prove the main theorems in this section, including the continuity of one-sided contagion value function (Theorem 2.4.2), viscosity solution representation (Theorem 2.6.2) and comparison principle (Theorem 2.6.6).

2.8.1   PROOF OF CONTINUITY OF ONE-SIDED CONTAGION VALUE FUNCTION

Proof. We prove the theorem in four steps: 1) v_0 is continuous in x, uniformly in t, s, p, 2) v_0 is continuous in s, uniformly in t, p, 3) v_0 is continuous in p, uniformly in t, s and 4) v_0 is continuous in t. Combining these four steps gives the continuity of v_0 in t, x, s, p.

Step 1. For any x_1, $x_2 \in [0, \infty)$ and $t, s, p \in [0, T] \times (0, \infty)^2$, using Assumption 2.3.2, we have

$$|v_0(t, x_1, s, p) - v_0(t, x_2, s, p)| = \left| \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^{t,x_1,s,p,\pi})\right] - \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^{t,x_2,s,p,\pi})\right] \right|$$

$$\leq \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[\left| U(X_T^{t,x_1,s,p,\pi}) - U(X_T^{t,x_2,s,p,\pi}) \right| \right]$$

$$\leq K \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[\left| X_T^{t,x_1,s,p,\pi} - X_T^{t,x_2,s,p,\pi} \right|^\gamma \right]$$

$$\leq K |x_1 - x_2|^\gamma,$$

by virtue of (2.2). Therefore, v_0 is continuous in x, uniformly in t, s, p.

Step 2. Fix $0 < s_1 < s_2 < \infty$ and $t, x, p \in [0, T] \times [0, \infty) \times (0, \infty)$. Denote by S^i the stock price that starts from s_i, $i = 1, 2$, and h^i and τ_i the corresponding default intensity and default time of stock P, respectively. By our model setting, τ_i can be represented by

$$\tau_i := \inf \left\{ s \geq t : \int_t^s h_u^i du \geq \mathcal{X} \right\},$$

where $\mathcal{X}$ is a standard exponential random variable on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and is independent of the filtration $(\mathcal{F}_t)_{t \geq 0}$, which means τ_i are totally inaccessible stopping times.

Define $\tau_{min} := \min\{\tau_1, \tau_2\}$. It is clear that before τ_{min}, the stock price dynamic is a standard geometric Brownian motion. We have

$$\mathbb{E}\left[\left| S_u^1 - S_u^2 \right| \mathbb{I}_{\{u < \tau_{min}\}} \right] \leq K |s_1 - s_2|$$

and

$$\mathbb{E}\left[\int_t^{\tau_{min} \wedge \overline{u}} \left| h_u^1 - h_u^2 \right| du \right] \leq K \mathbb{E}\left[\int_t^{\overline{u}} \left| S_u^1 - S_u^2 \right| \mathbb{I}_{\{u < \tau_{min}\}} du \right] \leq K |s_1 - s_2| \quad (2.19)$$

for any $\overline{u} \in [t, T]$.

If there is no jump on interval $[t, T]$, then $\sup_{[t,T]} |H_u^1 - H_u^2| = 0$ and $X_T^{t,x,s_1,p,\pi} = X_T^{t,x,s_2,p,\pi}$, where H^i is the jump process associated with default time τ_i. If there is

at least one jump on interval $[t, T]$, then $\sup_{[t,T]} |H_u^1 - H_u^2| = 1$ as τ_1 and τ_2 do not jump at the same time. We have the relation

$$
\begin{aligned}
|X_T^{t,x,s_1,p,\pi} - X_T^{t,x,s_2,p,\pi}| &= |X_T^{t,x,s_1,p,\pi} - X_T^{t,x,s_2,p,\pi}| \sup_{[t,T]} |H_u^1 - H_u^2| \\
&\leq (|X_T^{t,x,s_1,p,\pi}| + |X_T^{t,x,s_2,p,\pi}|) \sup_{[t,T]} |H_u^1 - H_u^2|.
\end{aligned}
$$

Since $\sup_{[t,T]} |H_u^1 - H_u^2|$ equals 0 or 1, we have $(\sup_{[t,T]} |H_u^1 - H_u^2|)^\alpha = \sup_{[t,T]} |H_u^1 - H_u^2|$ for any $\alpha > 0$. Using $(x + y)^\gamma \leq x^\gamma + y^\gamma$ for $x, y \geq 0$ and $0 < \gamma \leq 1$ and the Cauchy-Schwarz inequality, also noting Remark 2.3.1, we have

$$
\begin{aligned}
\mathbb{E}\left[|X_T^{t,x,s_1,p,\pi} - X_T^{t,x,s_2,p,\pi}|^\gamma\right] &\leq \mathbb{E}\left[(|X_T^{t,x,s_1,p,\pi}|^\gamma + |X_T^{t,x,s_2,p,\pi}|^\gamma) \sup_{[t,T]} |H_u^1 - H_u^2|\right] \\
&\leq K\left((\mathbb{E}\left[|X_T^{t,x,s_1,p,\pi}|^{2\gamma}\right])^{1/2} + (\mathbb{E}\left[|X_T^{t,x,s_2,p,\pi}|^{2\gamma}\right])^{1/2}\right) \\
&\quad \cdot \left(\mathbb{E}\left[\sup_{[t,T]} |H_u^1 - H_u^2|\right]\right)^{1/2} \\
&\leq Kx^\gamma \left(\mathbb{E}\left[\sup_{[t,T]} |H_u^1 - H_u^2|\right]\right)^{1/2}.
\end{aligned}
$$

We therefore have

$$
\begin{aligned}
|v_0(t, x, s_1, p) - v_0(t, x, s_2, p)| &\leq K \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[|X_T^{t,x,s_1,p,\pi} - X_T^{t,x,s_2,p,\pi}|^\gamma\right] \\
&\leq Kx^\gamma \left(\mathbb{E}\left[\sup_{[t,T]} |H_u^1 - H_u^2|\right]\right)^{1/2}.
\end{aligned}
$$

We can decompose H^i as $H_u^i = M_u^i + A_u^i$, where M^i is a martingale and $A_u^i := \int_t^{u \wedge \tau_i} h_s^i ds$ is a bounded variation process, see Bielecki and Rutkowski (2003). Ap-

plying Doob's sub-martingale inequality, we have

$$
\begin{aligned}
\mathbb{E}\left[\sup_{[t,T]}|H_u^1 - H_u^2|\right] &= \mathbb{E}\left[\sup_{[t,T]}|H_u^1 - H_u^2|^2\right] \\
&\leq 2\mathbb{E}\left[\sup_{[t,T]}|M_u^1 - M_u^2|^2 + \sup_{[t,T]}|A_u^1 - A_u^2|^2\right] \\
&\leq 8\mathbb{E}\left[|M_T^1 - M_T^2|^2\right] + 2\mathbb{E}\left[\sup_{[t,T]}|A_u^1 - A_u^2|^2\right].
\end{aligned}
$$

Since h is a monotone function of s by Assumption 2.2.1, without loss of generality, we assume h is non-increasing in s, then $h^1 \geq h^2$ before the first default occurs. By the definition of τ_i, we have $\tau_1 \leq \tau_2$ and $H_t^1 \geq H_t^2$. Then

$$
\begin{aligned}
\left|A_u^1 - A_u^2\right| &= \left|\int_t^u h_s^1 ds - \int_t^u h_s^2 ds\right|\mathbb{I}_{\{u \leq \tau_1 \leq \tau_2\}} + \left|\int_t^{\tau_1} h_s^1 ds - \int_t^u h_s^2 ds\right|\mathbb{I}_{\{\tau_1 < u \leq \tau_2\}} \\
&\quad + \left|\int_t^{\tau_1} h_s^1 ds - \int_t^{\tau_2} h_s^2 ds\right|\mathbb{I}_{\{\tau_1 \leq \tau_2 < u\}} \\
&= \left(\int_t^u h_s^1 ds - \int_t^u h_s^2 ds\right)\mathbb{I}_{\{u \leq \tau_1 \leq \tau_2\}} + \left(\mathcal{X} - \int_t^u h_s^2 ds\right)\mathbb{I}_{\{\tau_1 < u \leq \tau_2\}} \\
&\quad + (\mathcal{X} - \mathcal{X})\mathbb{I}_{\{\tau_1 \leq \tau_2 < u\}} \\
&= A_u^1 - A_u^2
\end{aligned}
$$

for any $u \in [t, T]$. Therefore,

$$
\sup_{[t,T]}|A_u^1 - A_u^2|^2 \leq K\sup_{[t,T]}|A_u^1 - A_u^2| = K\sup_{[t,T]}(A_u^1 - A_u^2).
$$

Note that $A_u^1 - A_u^2$ is non-decreasing before $\tau_1 \wedge T$ and non-increasing after $\tau_1 \wedge T$, we conclude that

$$
\sup_{[t,T]}(A_u^1 - A_u^2) = A_{\tau_1 \wedge T}^1 - A_{\tau_1 \wedge T}^2 = \int_t^{\tau_1 \wedge T}(h_u^1 - h_u^2)du.
$$

47

By inequality (2.19), we have

$$\mathbb{E}\left[\sup_{[t,T]}|A_u^1 - A_u^2|^2\right] \le K\mathbb{E}\left[\int_t^{\tau_1 \wedge T}(h_u^1 - h_u^2)du\right] \le K|s_1 - s_2|. \qquad (2.20)$$

Since $H_T^1 - H_T^2$ equals 0 or 1, we have

$$
\begin{aligned}
|M_T^1 - M_T^2|^2 &\le 2|H_T^1 - H_T^2|^2 + 2|A_T^1 - A_T^2|^2 \\
&\le 2(H_T^1 - H_T^2) + K(A_T^1 - A_T^2) \\
&\le 2(M_T^1 - M_T^2) + K(A_T^1 - A_T^2).
\end{aligned}
$$

Since M^i is martingale, also note that $\tau_1 \wedge T \le \tau_2 \wedge T$, we have

$$
\begin{aligned}
\mathbb{E}|M_T^1 - M_T^2|^2 &\le K\mathbb{E}[A_T^1 - A_T^2] \\
&= K\mathbb{E}\left[\int_t^{\tau_1 \wedge T} h_s^1 ds - \int_t^{\tau_2 \wedge T} h_s^2 ds\right] \\
&\le K\mathbb{E}\left[\int_t^{\tau_1 \wedge T}(h_s^1 - h_s^2)ds\right] \\
&\le K|s_1 - s_2|.
\end{aligned}
\qquad (2.21)
$$

Combining (2.21) and (2.20), we conclude that $\mathbb{E}\left[\sup_{[t,T]}|H_u^1 - H_u^2|\right] \le K|s_1 - s_2|$, which gives

$$|v_0(t, x, s_1, p) - v_0(t, x, s_2, p)| \le Kx^\gamma |s_1 - s_2|^{\frac{1}{2}}.$$

Therefore, v_0 is continuous in s, uniformly in t, p.

Step 3. Fix $0 < p_1 < p_2 < \infty$ and $t, x, s \in [0, T] \times [0, \infty) \times (0, \infty)$, by same technique as in Step 2, we can show

$$|v_0(t, x, s, p_1) - v_0(t, x, s, p_2)| \le Kx^\gamma |p_1 - p_2|^{\frac{1}{2}}.$$

Therefore, v_0 is continuous in p, uniformly in t, s.

Step 4. For any $0 \le t_1 < t_2 \le T$ and $x, s, p \in [0, \infty) \times (0, \infty)^2$, by the definition of

v_0 and the dynamic programming principle, $\forall \delta > 0$, $\exists \pi(\delta) \in \mathcal{A}$ such that

$$v_0(t_1, x, s, p) - \delta \leq \mathbb{E}\left[v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) \mathbb{I}_{\{t_2 < \tau\}} + v_1\left(t_2, X_{t_2}^{t_1,x,\pi(\delta)}\right) \mathbb{I}_{\{t_2 \geq \tau\}}\right]$$

$$\leq v_0(t_1, x, s, p).$$

Rearranging the order, we have

$$|v_0(t_1, x, s, p) - v_0(t_2, x, s, p)| - \delta$$

$$\leq \left| \mathbb{E}\left[v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) \mathbb{I}_{\{t_2 < \tau\}} + v_1\left(t_2, X_{t_2}^{t_1,x,\pi(\delta)}\right) \mathbb{I}_{\{t_2 \geq \tau\}}\right] - v_0(t_2, x, s, p) \right|$$

$$\leq \mathbb{E}\left[\left| v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) \mathbb{I}_{\{t_2 < \tau\}} - v_0(t_2, x, s, p)\right| \right]$$

$$+ \mathbb{E}\left[\left| v_1\left(t_2, X_{t_2}^{t_1,x,\pi(\delta)}\right) \mathbb{I}_{\{t_2 \geq \tau\}}\right| \right].$$

Using the Cauchy-Schwartz inequality, we have

$$\mathbb{E}\left[\left| v_1\left(t_2, X_{t_2}^{t_1,x,s,\pi(\delta)}\right) \mathbb{I}_{\{t_2 \geq \tau\}}\right| \right] \leq \mathbb{E}\left[\left| v_1\left(t_2, X_{t_2}^{t_1,x,s,\pi(\delta)}\right) \right|^2 \right]^{1/2} \sqrt{\mathbb{P}\left(t_2 \geq \tau\right)}$$

$$\leq K(1 + x^\gamma)\sqrt{\mathbb{P}\left(t_2 \geq \tau\right)},$$

which tends to 0 since $\mathbb{P}\left(t_2 \geq \tau\right) \to 0$ as $t_2 - t_1 \to 0$.

Next we prove the first term $\mathbb{E}\left[\left| v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) \mathbb{I}_{\{t_2 < \tau\}} - v_0(t_2, x, s, p)\right| \right]$ goes to zero as $t_2 - t_1 \to 0$.

$$\mathbb{E}\left[\left| v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) \mathbb{I}_{\{t_2 < \tau\}} - v_0(t_2, x, s, p)\right| \right]$$

$$\leq \mathbb{E}\left[\left| \left(v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) - v_0\left(t_2, x, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right)\right) \mathbb{I}_{\{t_2 < \tau\}}\right| \right]$$

$$+ \mathbb{E}\left[\left| \left(v_0\left(t_2, x, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right) - v_0(t_2, x, s, P_{t_2}^{t_1,p})\right) \mathbb{I}_{\{t_2 < \tau\}}\right| \right]$$

$$+ \mathbb{E}\left[\left| \left(v_0\left(t_2, x, s, P_{t_2}^{t_1,p}\right) - v_0(t_2, x, s, p)\right) \mathbb{I}_{\{t_2 < \tau\}}\right| \right] + \mathbb{E}\left[\left| v_0(t_2, x, s, p)\mathbb{I}_{\{t_2 \geq \tau\}}\right| \right].$$

As shown in Step 1,

$$\left| v_0 \left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p} \right) - v_0 \left(t_2, x, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p} \right) \right| \le K \left| X_{t_2}^{t_1,x,s,p,\pi(\delta)} - x \right|^\gamma,$$

and by (2.2),

$$\mathbb{E} \left[\left| X_{t_2}^{t_1,x,s,p,\pi(\delta)} - x \right|^2 \right] \le 2x^2 + 2\mathbb{E} \left[\left| X_{t_2}^{t_1,x,s,p,\pi(\delta)} \right|^2 \right] < \infty.$$

Therefore, $\left| X_{t_2}^{t_1,x,s,p,\pi(\delta)} - x \right|^\gamma$ is uniformly integrable, and we can exchange the order of expectation and limit to get

$$\lim_{t_2-t_1 \to 0} \mathbb{E} \left[\left| v_0 \left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p} \right) - v_0 \left(t_2, x, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p} \right) \right| \mathbb{I}_{\{t_2<\tau\}} \right]$$
$$\le \mathbb{E} \left[K \lim_{t_2-t_1 \to 0} \left| X_{t_2}^{t_1,x,s,p,\pi(\delta)} - x \right|^\gamma \right] = 0.$$

The same argument can be applied to the term

$$\mathbb{E} \left[\left| v_0 \left(t_2, x, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p} \right) - v_0(t_2, x, s, P_{t_2}^{t_1,p}) | \mathbb{I}_{\{t_2<\tau\}} \right| \right]$$

based on Step 2 and

$$\mathbb{E} \left[\left| v_0 \left(t_2, x, s, P_{t_2}^{t_1,p} \right) - v_0(t_2, x, s, p) | \mathbb{I}_{\{t_2<\tau\}} \right| \right]$$

based on Step 3, and we conclude that

$$\lim_{t_2-t_1 \to 0} \mathbb{E} \left[\left| v_0 \left(t_2, x, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p} \right) - v_0(t_2, x, s, P_{t_2}^{t_1,p}) | \mathbb{I}_{\{t_2<\tau\}} \right| \right]$$
$$\le \mathbb{E} \left[K x^\gamma \lim_{t_2-t_1 \to 0} \left| S_{t_2}^{t_1,s} - s \right|^{\frac{1}{2}} \right] = 0$$

and

$$\lim_{t_2-t_1\to 0} \mathbb{E}\left[\left|v_0\left(t_2, x, s, P_{t_2}^{t_1,p}\right) - v_0(t_2, x, s, p)|\mathbb{I}_{\{t_2<\tau\}}\right|\right]$$

$$\leq \mathbb{E}\left[Kx^\gamma \lim_{t_2-t_1\to 0} \left|P_{t_2}^{t_1,p} - p\right|^{\frac{1}{2}}\right] = 0.$$

The last term $|v_0(t_2, x, s, p)|\mathbb{P}(t_2 \geq \tau) \leq K(1 + x^\gamma)\mathbb{P}(t_2 \geq \tau)$, which tends to zero when $t_2 - t_1 \to 0$. Therefore

$$\mathbb{E}\left[\left|v_0\left(t_2, X_{t_2}^{t_1,x,s,p,\pi(\delta)}, S_{t_2}^{t_1,s}, P_{t_2}^{t_1,p}\right)\mathbb{I}_{\{t_2<\tau\}} - v_0(t_2, x, s, p)\right|\right] \to 0$$

as $t_2 - t_1 \to 0$ and we finally have

$$\lim_{t_2-t_1\to 0} |v_0(t_1, x, s, p) - v_0(t_2, x, s, p)| \leq \delta.$$

Since δ is arbitrary, we conclude that $v_0(t, x, s, p)$ is continuous in t. Combining Steps 1,2,3,4, we conclude that $v_0(t, x, s, p)$ is continuous in $[0, T] \times [0, \infty) \times (0, \infty)^2$. □

2.8.2 PROOF OF VISCOSITY SOLUTION

We first show the proof of the following lemma that is used in the viscosity subsolution proof.

Lemma 2.8.1. *Denote by $R_z := R_z(t, x, s)$ the following function*

$$R_z := \sup_{\pi\in A}\left\{\theta^T\pi x\frac{\partial w_z}{\partial x} + \frac{1}{2}\pi^T\Sigma\pi x^2\frac{\partial^2 w_z}{\partial x^2} + \sum_{i\in I_z}\rho_i^T\sigma\pi\sigma_i x s_i\frac{\partial^2 w_z}{\partial x\partial s_i}\right.$$

$$\left. + \sum_{i\in I_z}h_z^i(s)w_{z^i}\left(t, x\left(1 - \sum_{j=1}^N L_{ji}\pi^j\right), s^i\right)\right\},$$

where $w_z \in C^{1,2,\dots,2}$ for $\forall z \in I$. Then R_z is continuous in (t, x, s).

Proof. Let $z \in I$ and the point $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty)^{N_z+1}$ and $B_\eta(\bar{t}, \bar{x}, \bar{s})$ be the

ball with center $(\bar{t}, \bar{x}, \bar{s})$ and radius η. By the definition of supremum function, for any $\delta > 0$, there exists a control $\pi \in A$ such that

$$
\begin{aligned}
R_z(\bar{t}, \bar{x}, \bar{s}) - \delta \;\leq\; & \theta^T \pi \bar{x} \frac{\partial w_z}{\partial x}(\bar{t}, \bar{x}, \bar{s}) + \frac{1}{2} \pi^T \Sigma \pi \bar{x}^2 \frac{\partial^2 w_z}{\partial x^2}(\bar{t}, \bar{x}, \bar{s}) \\
& + \sum_{i \in I_z} \rho_i^T \sigma \pi \sigma_i \bar{x} \bar{s}_i \frac{\partial^2 w_z}{\partial x \partial s_i}(\bar{t}, \bar{x}, \bar{s}) \\
& + \sum_{i \in I_z} h_z^i(\bar{s}) w_{z^i}\left(\bar{t}, \bar{x}\left(1 - \sum_{j=1}^{N} L_{ji}\pi^j\right), \bar{s}^i\right),
\end{aligned}
\qquad (2.22)
$$

For any point $(t, x, s) \in B_\eta(\bar{t}, \bar{x}, \bar{s})$, we have

$$
\begin{aligned}
R_z(t, x, s) \;\geq\; & \theta^T \pi x \frac{\partial w_z}{\partial x}(t, x, s) + \frac{1}{2} \pi^T \Sigma \pi x^2 \frac{\partial^2 w_z}{\partial x^2}(t, x, s) \\
& + \sum_{i \in I_z} \rho_i^T \sigma \pi \sigma_i x s_i \frac{\partial^2 w_z}{\partial x \partial s_i}(t, x, s) \\
& + \sum_{i \in I_z} h_z^i(s) w_{z^i}\left(t, x\left(1 - \sum_{j=1}^{N} L_{ji}\pi^j\right), s^i\right),
\end{aligned}
\qquad (2.23)
$$

Subtracting (2.23) from (2.22), we have

$$
\begin{aligned}
& R_z(\bar{t}, \bar{x}, \bar{s}) - R_z(t, x, s) - \delta \\
\leq\; & \theta^T \pi \left(\bar{x} \frac{\partial w_z}{\partial x}(\bar{t}, \bar{x}, \bar{s}) - x \frac{\partial w_z}{\partial x}(t, x, s)\right) + \frac{1}{2} \pi^T \Sigma \pi \left(\bar{x}^2 \frac{\partial^2 w_z}{\partial x^2}(\bar{t}, \bar{x}, \bar{s}) - x^2 \frac{\partial^2 w_z}{\partial x^2}(t, x, s)\right) \\
& + \sum_{i \in I_z} \rho_i^T \sigma \pi \sigma_i \left(\bar{x} \bar{s}_i \frac{\partial^2 w_z}{\partial x \partial s_i}(\bar{t}, \bar{x}, \bar{s}) - x s_i \frac{\partial^2 w_z}{\partial x \partial s_i}(t, x, s)\right) \\
& + \sum_{i \in I_z} \left(h_z^i(\bar{s}) w_{z^i}\left(\bar{t}, \bar{x}\left(1 - \sum_{j=1}^{N} L_{ji}\pi^j\right), \bar{s}^i\right) - h_z^i(s) w_{z^i}\left(t, x\left(1 - \sum_{j=1}^{N} L_{ji}\pi^j\right), s^i\right)\right).
\end{aligned}
$$

Taking the limit superior and then letting δ tend to 0, we have

$$
R_z(\bar{t}, \bar{x}, \bar{s}) \leq \liminf_{(t,x,s) \to (\bar{t}, \bar{x}, \bar{s})} R_z(t, x, s).
\qquad (2.24)
$$

Similarly, using the smoothness of w_z and boundedness of π and h_z^i for $\forall z \in I$ and

52

$i \in \{1, \ldots, N\}$, we can get

$$R_z(\bar{t}, \bar{x}, \bar{s}) \geq \limsup_{(t,x,s) \to (\bar{t}, \bar{x}, \bar{s})} R_z(t, x, s). \qquad (2.25)$$

Inequalities (2.24) and (2.25) imply that $R_z(t, x, s)$ is continuous in (t, x, s) for $\forall z \in I$. $\qquad \square$

Combining Lemma 2.8.1 and the smoothness of w, we conclude that F_z given in Definition 2.6.1 is continuous in (t, x, s). Based on this result, we prove the value function $v := (v_z)_{z \in I}$ is a viscosity solution to the PDE system (2.3).

Proposition 2.8.2. *The value function $v := (v_z)_{z \in I}$ is a viscosity supersolution to equation (2.3) on $[0, T) \times (0, \infty)^{N+1}$.*

Proof. Let $\bar{z} \in I$, $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty)^{N_{\bar{z}}+1}$ and $\varphi := (\varphi_z)_{z \in I}$ be tuple of test functions where $\varphi_z \in C^{1,2,\ldots,2}\left([0, T) \times (0, \infty)^{N_z+1}\right)$ for $\forall z \in I$ such that

$$0 = ((v_{\bar{z}})_* - \varphi_{\bar{z}})(\bar{t}, \bar{x}, \bar{s}) = \min_{[0,T) \times (0,\infty)^{N_{\bar{z}}+1}} ((v_{\bar{z}})_* - \varphi_{\bar{z}})(t, x, s), \qquad (2.26)$$

and $(v_z)_* \geq \varphi_z$ for $\forall z \in I$ on $[0, T) \times (0, \infty)^{N_z+1}$.

By definition of $(v_{\bar{z}})_*$, there exists a sequence (t_m, x_m, s_m) in $[0, T) \times (0, \infty)^{N_{\bar{z}}+1}$ such that

$$(t_m, x_m, s_m) \to (\bar{t}, \bar{x}, \bar{s}) \quad \text{and} \quad v_{\bar{z}}(t_m, x_m, s_m) \to (v_{\bar{z}})_*(\bar{t}, \bar{x}, \bar{s}),$$

when m goes to infinity. By the continuity of $\varphi_{\bar{z}}$ and by (2.26) we also have that

$$\gamma_m := v_{\bar{z}}(t_m, x_m, s_m) - \varphi_{\bar{z}}(t_m, x_m, s_m) \to 0,$$

when m goes to infinity.

Let $\pi \in \mathcal{A}$ be a constant control process and $B_\eta(x_m, s_m) \in (0, \infty)^{N_{\bar{z}}+1}$ be the ball with center (x_m, s_m) and radius $\eta > 0$. Note that when m is large enough,

$(x_m, s_m) \in B_\eta(\bar{x}, \bar{s})$, thus $\forall (x, s) \in B_\eta(x_m, s_m)$, we have $(x, s) \in B_{2\eta}(\bar{x}, \bar{s})$. We denote by $X_u^{t_m, x_m}$ the associated controlled wealth process. Let τ_m^π be the stopping time given by

$$\tau_m^\pi := \inf \left\{ u \in [t_m, T) : \left(X_u^{t_m, x_m}, S_u^{t_m, s_m} \right) \notin B_\eta(x_m, s_m) \right\}.$$

Let (h_m) be a strictly positive sequence such that

$$h_m \to 0 \quad \text{and} \quad \frac{\gamma_m}{h_m} \to 0,$$

when m goes to infinity. Then we can define a stopping time θ_m give by $\theta_m := \tau_m^\pi \wedge (t_m + h_m) \wedge \widetilde{\tau}_m$ where $\widetilde{\tau}_m$ is the first default time of the surviving stocks, starting from t_m.

Next we use the weak dynamic programming principle (weak-DPP) proved in Bouchard and Touzi (2011), that is,

$$v_{\bar{z}}(t, x, s) \geq \mathbb{E}\left[\sum_{z \in I} (v_z)_* \left(\theta, X_\theta^{t, x, s, \bar{z}}, S_\theta^{t, s} \right) \mathbb{I}_{\{\mathbb{H}_\theta = z\}} \right],$$

for any $\mathcal{G}$-measurable stopping time $\theta \in [t, T]$ such that X_θ and S_θ are $\mathbb{L}^\infty$–bounded.

Since under stopping time θ_m, the processes S and X are both bounded, we can apply above weak dynamic programming principle (weak-DPP) for $v_{\bar{z}}(t_m, x_m, s_m)$ to θ_m and get

$$v_{\bar{z}}(t_m, x_m, s_m) \geq \mathbb{E}\left[\sum_{z \in I} (v_z)_* \left(\theta_m, X_{\theta_m}^{t_m, x_m, s_m, \bar{z}}, S_{\theta_m}^{t_m, s_m} \right) \mathbb{I}_{\{\mathbb{H}_{\theta_m} = z\}} \right].$$

Equation (2.26) implies $(v_z)_* \geq \varphi_z$ for $\forall z \in I$, thus

$$\varphi_{\bar{z}}(t_m, x_m, s_m) + \gamma_m \geq \mathbb{E}\left[\sum_{z \in I} \varphi_z \left(\theta_m, X_{\theta_m}^{t_m, x_m, s_m, \bar{z}}, S_{\theta_m}^{t_m, s_m} \right) \mathbb{I}_{\{\mathbb{H}_{\theta_m} = z\}} \right].$$

Applying Ito's formula to the whole term in bracket, we obtain

$$\frac{\gamma_m}{h_m} - \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{\theta_m}\sum_{z\in I}\mathcal{L}^\pi\varphi_z(u, X_u^\pi, S_u)\mathbb{I}_{\{\mathbb{H}_u=z\}}du\right] \geq 0 \qquad (2.27)$$

after noting that the stochastic integral term cancels out by taking expectations since the integrand is bounded. Note that $\mathcal{L}^\pi\varphi_z\left(u, X_u^{t_m,x_m}, S_u^{t_m,s_m}\right)$ is defined the same as (2.4).

Next we investigate the stopping time θ_m when m is large enough. Firstly, for the stopping time τ_m^π, denoting $E_m := \{\widetilde{\tau}_m > t_m + h_m\}$, we have

$$\begin{aligned}
&\mathbb{P}\left(\tau_m^\pi \leq t_m + h_m \mid E_m\right) \\
=\ &\mathbb{P}\left(\sup_{t\in[t_m,t_m+h_m]}\left(\left|X_t^{t_m,x_m} - x_m\right|^2 + \sum_{i\in I_{\bar{z}}}\left|S_t^{i,t_m,s_m} - s_{m,i}\right|^2\right) \geq \eta^2 \mid E_m\right) \\
\leq\ &\mathbb{P}\left(\sup_{t\in[t_m,t_m+h_m]}\left|X_t^{t_m,x_m} - x_m\right|^2 \geq \frac{\eta^2}{N_{\bar{z}}+1} \mid E_m\right) \\
&+ \sum_{i\in I_{\bar{z}}}\mathbb{P}\left(\sup_{t\in[t_m,t_m+h_m]}\left|S_t^{i,t_m,s_m} - s_{m,i}\right|^2 \geq \frac{\eta^2}{N_{\bar{z}}+1} \mid E_m\right) \\
\leq\ &\frac{N_{\bar{z}}+1}{\eta^2}\left(\mathbb{E}\left[\sup_{t\in[t_m,t_m+h_m]}\left|X_t^{t_m,x_m} - x_m\right|^2 \mid E_m\right]\right. \\
&\left.+ \sum_{i\in I_{\bar{z}}}\mathbb{E}\left[\sup_{t\in[t_m,t_m+h_m]}\left|S_t^{i,t_m,s_m} - s_{m,i}\right|^2 \mid E_m\right]\right).
\end{aligned}$$

By Pham (2009), Page 67, each term in the bracket converges to zero as $m \to \infty$, which gives

$$\lim_{m\to\infty}\mathbb{P}\left(\tau_m^\pi \leq t_m + h_m \mid E_m\right) = 0.$$

By definition of conversion time $\widetilde{\tau}_m$, we have

$$\mathbb{P}\left(E_m^C\right) = 1 - \mathbb{E}\left[e^{-\int_{t_m}^{t_m+h_m}\sum_{i\in I_{\bar{z}}}h_{\bar{z}}^i(S_u)du}\right] \leq 1 - e^{-Kh_m}$$

due to the boundedness of intensity function $h^i_{\bar{z}}$. Thus

$$\lim_{m\to\infty} \mathbb{P}(E^C_m) = 0.$$

Finally for the stopping time $\tau^\pi_m \wedge \widetilde{\tau}_m$, we have

$$
\begin{aligned}
\mathbb{P}\left(\tau^\pi_m \wedge \widetilde{\tau}_m \leq t_m + h_m\right) & \leq & \mathbb{P}\left(\tau^\pi_m \leq t_m + h_m\right) + \mathbb{P}\left(E^C_m\right) \\
& = & \mathbb{P}\left(\tau^\pi_m \leq t_m + h_m, E_m\right) + \mathbb{P}\left(\tau^\pi_m \leq t_m + h_m, E^C_m\right) + \mathbb{P}\left(E^C_m\right) \\
& \leq & \mathbb{P}\left(\tau^\pi_m \leq t_m + h_m \mid E_m\right) + 2\mathbb{P}\left(E^C_m\right). \quad (2.28)
\end{aligned}
$$

Combining above results, we get

$$\lim_{m\to\infty} \mathbb{P}\left(\tau^\pi_m \wedge \widetilde{\tau}_m \leq t_m + h_m\right) = 0.$$

We now estimate

$$
\begin{aligned}
-\frac{\gamma_m}{h_m} & \leq & \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{\theta_m} -\sum_{z\in I}\mathcal{L}^\pi\varphi_z(u, X^\pi_u, S_u)\mathbb{I}_{\{\mathbb{H}_u=z\}}du\right] \\
& \leq & \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{t_m+h_m} -\sum_{z\in I}\mathcal{L}^\pi\varphi_z(u, X^\pi_u, S_u)\mathbb{I}_{\{\mathbb{H}_u=z\}}du \mid \tau^\pi_m \wedge \widetilde{\tau}_m > t_m + h_m\right] \\
& & \cdot \mathbb{P}\left(\tau^\pi_m \wedge \widetilde{\tau}_m > t_m + h_m\right) \\
& & + \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{\tau^\pi_m\wedge\widetilde{\tau}_m} -\sum_{z\in I}\mathcal{L}^\pi\varphi_z(u, X^\pi_u, S_u)\mathbb{I}_{\{\mathbb{H}_u=z\}}du \mid \tau^\pi_m \wedge \widetilde{\tau}_m \leq t_m + h_m\right] \\
& & \cdot \mathbb{P}\left(\tau^\pi_m \wedge \widetilde{\tau}_m \leq t_m + h_m\right) \\
& \leq & \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{t_m+h_m} -\mathcal{L}^\pi\varphi_{\bar{z}}\left(u, X^{t_m,x_m}_u, S^{t_m,s_m}_u\right)du \mid \tau^\pi_m \wedge \widetilde{\tau}_m > t_m + h_m\right] \\
& & + K\mathbb{P}\left(\tau^\pi_m \wedge \widetilde{\tau}_m \leq t_m + h_m\right).
\end{aligned}
$$

By the mean value theorem and dominated convergence theorem, taking limit on both sides of the inequality, we have

$$-\mathcal{L}^\pi\varphi_{\bar{z}}(\bar{t}, \bar{x}, \bar{s}) \geq 0,$$

which implies

$$F_{\bar{z}}\left(\bar{t}, \bar{x}, \bar{s}, \varphi, \nabla_{(t,x,s)}\varphi_{\bar{z}}, \nabla^2_{(x,s)}\varphi_{\bar{z}}\right) \geq 0,$$

due to the arbitrariness of $\pi \in A$. $\qquad\qquad\square$

Proposition 2.8.3. *The value function* $v := (v_z)_{z \in I}$ *is a viscosity subsolution to equation (2.3) on* $[0,T) \times (0, \infty)^{N+1}$.

Proof. Let $\bar{z} \in I$, $(\bar{t}, \bar{x}, \bar{s}) \in [0,T) \times (0, \infty)^{N_{\bar{z}}+1}$ and $\varphi := (\varphi_z)_{z \in I}$ be tuple of test functions where $\varphi_z \in C^{1,2,\ldots,2}\left([0,T) \times (0, \infty)^{N_z+1}\right)$ for $\forall z \in I$ such that

$$0 = \left((v_{\bar{z}})^* - \varphi_{\bar{z}}\right)(\bar{t}, \bar{x}, \bar{s}) = \max_{[0,T) \times (0,\infty)^{N_{\bar{z}}+1}} \left((v_{\bar{z}})^* - \varphi_{\bar{z}}\right)(t, x, s), \qquad (2.29)$$

and $(v_z)^* \leq \varphi_z$ for $\forall z \in I$ on $[0,T) \times (0, \infty)^{N_z+1}$.

We prove the result by contradiction. Assume on the contrary that

$$F_{\bar{z}}\left(\bar{t}, \bar{x}, \bar{s}, \varphi, \nabla_{(t,x,s)}\varphi_{\bar{z}}, \nabla^2_{(x,s)}\varphi_{\bar{z}}\right) > 0$$

Then by the continuity of $F_{\bar{z}}$, there exists $\delta > 0$ and $\eta > 0$ such that

$$F_{\bar{z}}\left(t, x, s, \varphi, \nabla_{(t,x,s)}\varphi, \nabla^2_{(x,s)}\varphi\right) = -\sup_{\pi \in A} \mathcal{L}^{\pi}\varphi_{\bar{z}}(t, x, s) > \delta$$

for $(t, x, s) \in B_{\eta}(\bar{t}, \bar{x}, \bar{s})$. By definition of $(v_{\bar{z}})^*$, there exists a sequence (t_m, x_m, s_m) taking values in $B_{\frac{\eta}{2}}(\bar{t}, \bar{x}, \bar{s})$ such that

$$(t_m, x_m, s_m) \to (\bar{t}, \bar{x}, \bar{s}) \quad \text{and} \quad v_{\bar{z}}(t_m, x_m, s_m) \to (v_{\bar{z}})^*(\bar{t}, \bar{x}, \bar{s}),$$

when m goes to infinity. By the continuity of $\varphi_{\bar{z}}$ and by (2.29) we also have that

$$\gamma_m := v_{\bar{z}}(t_m, x_m, s_m) - \varphi_{\bar{z}}(t_m, x_m, s_m) \to 0,$$

when m goes to infinity.

We denote by $X_u^{t_m,x_m,\pi}$ the controlled wealth process associated with control process $\pi \in \mathcal{A}$. Let $\tau_m^{\pi^m}$ be the stopping time given by

$$\tau_m^\pi := \inf \left\{ u \in [t_m, T) : \left(u, X_u^{t_m,x_m,\pi}, S_u^{t_m,s_m} \right) \notin B_{\frac{\eta}{2}}(t_m, x_m, s_m) \right\}.$$

Let (h_m) be a strictly positive sequence such that

$$h_m \to 0 \quad \text{and} \quad \frac{\gamma_m}{h_m} \to 0,$$

when m goes to infinity. Then we can define a stopping time θ_m give by $\theta_m := \tau_m^\pi \wedge (t_m + h_m) \wedge \widetilde{\tau}_m$ where $\widetilde{\tau}_m$ is the first default time of surviving stocks starting from t_m.

Next we use the weak dynamic programming principle (weak-DPP) proved in Bouchard and Touzi (2011), that is, for any $\epsilon > 0$, there exists a control process $\pi \in \mathcal{A}$ such that

$$v_{\bar{z}}(t, x, s) - \epsilon \leq \mathbb{E}\left[\sum_{z \in I} (v_z)^* \left(\theta, X_\theta^{t,x,s,\bar{z},\pi}, S_\theta^{t,s} \right) \mathbb{I}_{\{\mathbb{H}_\theta = z\}} \right],$$

for any $\mathcal{G}$–stopping time $\theta \in [t, T]$.

We apply above weak dynamic programming principle (weak-DPP) for $v_{\bar{z}}(t_m, x_m, s_m)$ to θ_m and get for $\epsilon = \delta \frac{h_m}{2} > 0$, there exists $\pi \in \mathcal{A}$ such that

$$v_{\bar{z}}(t_m, x_m, s_m) - \delta \frac{h_m}{2} \leq \mathbb{E}\left[\sum_{z \in I} (v_z)^* \left(\theta_m, X_{\theta_m}^{t_m,x_m,s_m,\bar{z},\pi}, S_{\theta_m}^{t_m,s_m} \right) \mathbb{I}_{\{\mathbb{H}_{\theta_m} = z\}} \right].$$

Equation (2.29) implies $(v_z)^* \leq \varphi_z$ for $\forall z \in I$, thus

$$\varphi_{\bar{z}}(t_m, x_m, s_m) + \gamma_m - \delta \frac{h_m}{2} \leq \mathbb{E}\left[\sum_{z \in I} \varphi_z \left(\theta_m, X_{\theta_m}^{t_m,x_m,s_m,\bar{z},\pi}, S_{\theta_m}^{t_m,s_m} \right) \mathbb{I}_{\{\mathbb{H}_{\theta_m} = z\}} \right].$$

Applying Ito's formula to the whole term in bracket, we obtain

$$
\begin{aligned}
\frac{\gamma_m}{h_m} - \frac{\delta}{2} \;\le\;& \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{\theta_m}\sum_{z\in I}\mathcal{L}^\pi\varphi_z(u, X_u^\pi, S_u)\mathbb{I}_{\{\mathbb{H}_u=z\}}du\right] \\
\le\;& \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{t_m+h_m}\mathcal{L}^\pi\varphi_{\bar{z}}(u, X_u^\pi, S_u)du \mid \tau_m^\pi\wedge\widetilde{\tau}_m > t_m+h_m\right] \\
& + K\mathbb{P}\left(\tau_m^\pi\wedge\widetilde{\tau}_m \le t_m+h_m\right)
\end{aligned}
$$

after noting that the stochastic integral term cancels out by taking expectations since the integrand is bounded. By the similar technique as the supersolution proof, we can show that $\mathbb{P}\left(\tau_m^\pi\wedge\widetilde{\tau}_m \le t_m+h_m\right) \to 0$ as $m \to \infty$.

Since $\left(u, X_u^{t_m, x_m, \pi^m}, S_u^{t_m, s_m}\right) \in B_\eta(\bar{t}, \bar{x}, \bar{s})$ in $[t_m, t_m+h_m]$ if $\tau_m^\pi\wedge\widetilde{\tau}_m > t_m+h_m$, we have

$$
\mathcal{L}^\pi\varphi_{\bar{z}}\left(u, X_u^{t_m, x_m, \pi^m}, S_u^{t_m, s_m}\right) < -\delta
$$

in $[t_m, t_m+h_m]$. Thus

$$
\frac{\gamma_m}{h_m} - \frac{\delta}{2} \le \mathbb{E}\left[\frac{1}{h_m}\int_{t_m}^{t_m+h_m}-\delta du\right] + K\mathbb{P}\left(\tau_m^\pi\wedge\widetilde{\tau}_m \le t_m+h_m\right).
$$

Then we obtain

$$
\lim_{m\to\infty}\frac{\gamma_m}{h_m} - \frac{\delta}{2} \le -\delta,
$$

which implies $0 \le -\frac{\delta}{2}$. We thus get the desired contradiction with $\delta > 0$. $\qquad\square$

2.8.3 PROOF OF COMPARISON PRINCIPLE

To prove the comparison principle, we need an alternative definition of viscosity solution in terms of the notions of semijets defined as below.

Definition 2.8.4. For $z \in I$, given w_z a function on $[0, T) \times (0, \infty)^{N_z+1}$, the superjet

of w_z at $(t, x, s) \in [0, T) \times (0, \infty)^{N_z+1}$ is defined by:

$$\mathcal{P}^{1,2,\dots,2,+} w_z(t, x, s) = \Big\{ (R, q, Q) \in \mathbb{R} \times \mathbb{R}^{N_z+1} \times \mathcal{S}^{(N_z+1) \times (N_z+1)} \text{ such that}$$

$$w_z(t', x', s') \leq w_z(t, x, s) + R(t' - t) + \langle q, X' - X \rangle$$

$$+ \frac{1}{2} \langle Q(X' - X), X' - X \rangle + o(|t' - t|^2 + |X' - X|^2) \Big\},$$

where $X = (x, s)$, $X' = (x', s')$, and the bracket $\langle \cdot, \cdot \rangle$ is the inner product of two vectors. We define its closure $\bar{\mathcal{P}}^{1,2,\dots,2,+} w_z(t, x, s)$ as the set of elements $(R, q, Q) \in \mathbb{R} \times \mathbb{R}^{N_z+1} \times \mathcal{S}^{(N_z+1) \times (N_z+1)}$ for which there exists a sequence $(t_m, X_m, R_m, q_m, P_m) \in [0, T) \times (0, \infty)^{N_z+1} \times \mathcal{P}^{1,2,\dots,2,+} w_z(t, X)$ satisfying $(t_m, X_m, R_m, q_m, Q_m) \to (t, X, R, q, Q)$. We also define the subjets

$$\mathcal{P}^{1,2,\dots,2,-} w_z(t, x, s) = -\mathcal{P}^{1,2,\dots,2,+}(-w_z)(t, x, s),$$

$$\bar{\mathcal{P}}^{1,2,\dots,2,-} w_z(t, x, s) = -\bar{\mathcal{P}}^{1,2,\dots,2,+}(-w_z)(t, x, s).$$

By standard arguments, one has an equivalent definition of viscosity solutions in terms of semijets: $w := (w_z)_{z \in I}$ is a viscosity subsolution (resp. supersolution) to (2.3) at $(t, x, s) \in [0, T) \times (0, \infty)^{N+1}$ if and only if for all $z \in I$ and $(R, q, Q) \in \bar{\mathcal{P}}^{1,2,\dots,2,+} w_z(t, x, s)$ (resp. $\bar{\mathcal{P}}^{1,2,\dots,2,-} w_z(t, x, s)$).

$$F_z(t, x, s, w, (R, q), Q) \leq (resp. \geq) \, 0.$$

We can now state and prove the following comparison principle which gives rise to the uniqueness of viscosity solution.

Proposition 2.8.5. *Let $W := (W_z)_{z \in I}$ (resp. $V := (V_z)_{z \in I}$) be a u.s.c. viscosity subsolution (resp. l.s.c. viscosity supersolution) of (2.3) on $[0, T) \times (0, \infty)^{N+1}$ and satisfy the growth condition $|W_z(t, x, s)|, |V_z(t, x, s)| \leq K(1 + x^\gamma)$, the terminal relation $W_z(T, x, s) \leq V_z(T, x, s)$, and the boundary relations $W_z(t, x, s) \leq V_z(t, x, s)$ on the boundary of $[0, \infty)^{N_z+1}$ for $\forall z \in I$. Then we have $W_z \leq V_z$ for $\forall z \in I$ on*

$[0, T] \times [0, \infty)^{N_z+1}$.

Proof. We prove the result in several steps.

Step 1. Let $\widetilde{W}_z = e^{\Gamma t} W_z$ and $\widetilde{V}_z = e^{\Gamma t} V_z$ for constant $\Gamma > 0$, then a straightforward calculation shows that $\widetilde{W}$ (resp. $\widetilde{V}$) is a subsolution (resp. supersolution) of

$$-\sup_{\pi \in A} \widetilde{\mathcal{L}}^\pi w_z(t, x, s) = 0, \quad \text{on} \quad [0, T) \times (0, \infty)^{N_z+1},$$

for $z \in I$, where $\widetilde{\mathcal{L}}^\pi$ is given by

$$
\begin{aligned}
\widetilde{\mathcal{L}}^\pi w_z(t, x, s) = \ & \frac{\partial w_z}{\partial t} + (r + \theta^T \pi) x \frac{\partial w_z}{\partial x} + \sum_{i \in I_z} \mu_i s_i \frac{\partial w_z}{\partial s_i} \\
& + \frac{1}{2} \pi^T \Sigma \pi x^2 \frac{\partial^2 w_z}{\partial x^2} + \frac{1}{2} \sum_{i \in I_z} \sigma_i^2 s_i^2 \frac{\partial^2 w_z}{\partial s_i^2} \\
& + \sum_{i,j \in I_z, i<j} \rho_{ij} \sigma_i \sigma_j s_i s_j \frac{\partial^2 w_z}{\partial s_i \partial s_j} + \sum_{i \in I_z} \rho_i^T \sigma \pi \sigma_i x s_i \frac{\partial^2 w_z}{\partial x \partial s_i} \\
& + \sum_{i \in I_z} h_z^i(s) \left(w_{z^i} \left(t, x \left(1 - \sum_{j=1}^N L_{ji} \pi^j \right), s^i \right) - w_z \right) \\
& - \Gamma w_z.
\end{aligned}
\tag{2.30}
$$

We will show that $\widetilde{W}_z \leq \widetilde{V}_z$ for $\forall z \in I$ on $[0, T] \times [0, \infty)^{N_z+1}$ in the next few steps, thus we conclude $W_z \leq V_z$. We further define $\widetilde{F}$ function by

$$\widetilde{F}_z \left(t, x, s, w, \nabla_{(t,x,s)} w_z, \nabla^2_{(x,s)} w_z \right) = -\sup_{\pi \in A} \widetilde{\mathcal{L}}^\pi w_z(t, x, s).$$

Step 2. Define $\widetilde{V}_z^n := \widetilde{V}_z + \frac{1}{n} \phi_z(t, x, s)$, where $\phi_z(t, x, s) = e^{-\lambda t} \left(1 + x^{2\gamma} + \sum_{i \in I_z} s_i^{2\gamma} \right)$. We claim that $\widetilde{V}^n$ is a viscosity supersolution to (2.30). Note that

$$\mathcal{P}^{1,2,\dots,2,-} \widetilde{V}_z^n(t, x, s) = \mathcal{P}^{1,2,\dots,2,-} \widetilde{V}_z(t, x, s) + \frac{1}{n} (R', q', Q'),$$

where $R' = -\lambda\phi_z$, $q' = 2\gamma e^{-\lambda t}\left(x^{2\gamma-1}, s_1^{2\gamma-1}, \ldots, s_i^{2\gamma-1}, \ldots, s_N^{2\gamma-1}\right)$ for $i \in I_z$ and

$$
Q' = 2\gamma(2\gamma-1)e^{-\lambda t}
\begin{pmatrix}
x^{2\gamma-2} & 0 & \cdots & 0 \\
0 & s_1^{2\gamma-2} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & s_N^{2\gamma-2}
\end{pmatrix}.
$$

We have that for all $(R, q, Q) \in \mathcal{P}^{1,2,\ldots,2,-}\widetilde{V}_z^n(t, x, s)$,

$$
\left(R - \frac{R'}{n}, q - \frac{q'}{n}, Q - \frac{Q'}{n}\right) \in \mathcal{P}^{1,2,\ldots,2,-}\widetilde{V}_z(t, x, s).
$$

Since $\widetilde{V}$ is a viscosity supersolution to (2.30), we have

$$
\widetilde{F}_z\left(t, x, s, \widetilde{V}, \left(R - \frac{R'}{n}, q - \frac{q'}{n}\right), Q - \frac{Q'}{n}\right) \geq 0
$$

for $\forall z \in I$ by the equivalent definition of viscosity supersolution. Using the inequality $\sup\{A - B\} \geq \sup\{A\} - \sup\{B\}$ and the boundedness of controls and coefficients, we have

$$
\begin{aligned}
\widetilde{F}_z\left(t, x, s, \widetilde{V}^n, (R, q), Q\right) \;\geq\;& \widetilde{F}_z\left(t, x, s, \widetilde{V}, \left(R - \frac{R'}{n}, q - \frac{q'}{n}\right), Q - \frac{Q'}{n}\right) \\
& + \frac{1}{n}\left(\lambda + \Gamma + \sum_{i \in I_z} h_z^i(s) - K\right)\phi_z
\end{aligned}
$$

for a constant $K > 0$. Therefore, $\widetilde{F}_z\left(t, x, s, \widetilde{V}^n, (R, q), Q\right) \geq 0$ for a large enough λ, which implies that $\widetilde{V}^n$ is a viscosity supersolution of (2.30).

Step 3. We show that for all $n \geq 1$, it is $\widetilde{W}_z \leq \widetilde{V}_z^n$ for $z \in I$ on $[0, T) \times (0, \infty)^{N_z+1}$, and thus conclude that $\widetilde{W} \leq \widetilde{V}$. Fix $n \geq 1$ and define

$$
M_z := \sup_{X \in [0,T) \times (0,\infty)^{N_z+1}} [\widetilde{W}_z(X) - \widetilde{V}_z^n(X)],
$$

and

$$M := \max_{z \in I} M_z = M_{\bar{z}},$$

where $X := (t, x, s)$. We next show that $M \leq 0$. Suppose on the contrary that $M > 0$, by the growth condition on $\widetilde{W}_{\bar{z}}$ and $\widetilde{V}_{\bar{z}}$ we have

$$\lim_{x,s \to \infty} (\widetilde{W}_{\bar{z}} - \widetilde{V}_{\bar{z}}^n)(t, x, s) = -\infty$$

for any $t \in [0, T)$. By the terminal and boundary conditions, we also have

$$(\widetilde{W}_{\bar{z}} - \widetilde{V}_{\bar{z}}^n)(T, x, s) \leq 0, \ (\widetilde{W}_{\bar{z}} - \widetilde{V}_{\bar{z}}^n)(t, 0, s) \leq 0, \ (\widetilde{W}_{\bar{z}} - \widetilde{V}_{\bar{z}}^n)(t, x, 0) \leq 0.$$

Note that here $s = 0$ denotes $s_i = 0$ for any $i \in I_z$.

Since $\widetilde{W}_{\bar{z}} - \widetilde{V}_{\bar{z}}^n$ is upper-semicontinuous and $M > 0$, there exists some open bounded set $O \in [0, T) \times (0, \infty)^{N_{\bar{z}}+1}$ such that

$$M = \max_{X \in O}[\widetilde{W}_{\bar{z}}(X) - \widetilde{V}_{\bar{z}}^n(X)] > 0.$$

We now use the doubling variable technique. For any fixed $\epsilon > 0$, define

$$\Phi(X, Y) := \Phi_\epsilon(X, Y) = \widetilde{W}_{\bar{z}}(X) - \widetilde{V}_{\bar{z}}^n(Y) - \phi_1(X, Y),$$

where $\phi_1(X, Y) := \frac{1}{\epsilon}\|X - Y\|^2$. Note that Φ is upper-semicontinuous and hence achieves its maximum $\widetilde{M} = \widetilde{M}_\epsilon$ on the compact set $\bar{O}^2$ at $(\widetilde{X}, \widetilde{Y}) = (\widetilde{X}_\epsilon, \widetilde{Y}_\epsilon)$. We may write that, for all $\epsilon > 0$,

$$M \leq \widetilde{M} = \widetilde{W}_{\bar{z}}(\widetilde{X}) - \widetilde{V}_{\bar{z}}^n(\widetilde{Y}) - \phi_1(\widetilde{X}, \widetilde{Y}) \leq \widetilde{W}_{\bar{z}}(\widetilde{X}) - \widetilde{V}_{\bar{z}}^n(\widetilde{Y}).$$

The sequence $(\widetilde{X}, \widetilde{Y})$ converges, up to a subsequence, to some $(\hat{X}, \hat{Y}) \in \bar{O}^2$. Moreover, since $W_{\bar{z}}(\widetilde{X}) - V_{\bar{z}}^n(\widetilde{Y})$ is upper bounded due to the upper-semicontinuity of

$\widetilde{W}_{\bar{z}}$ and $-\widetilde{V}_{\bar{z}}^n$, we know $\phi_1(\widetilde{X}, \widetilde{Y})$ is bounded, which implies $\hat{X} = \hat{Y}$. Let ϵ tend to 0 and take the $\limsup$, we get $M \leq \widetilde{W}_{\bar{z}}(\hat{X}) - \widetilde{V}_{\bar{z}}^n(\hat{Y}) \leq M$. Therefore, $\hat{X} = \hat{Y} \in O$ and $\phi_1(\widetilde{X}, \widetilde{Y}) \to 0$.

Step 4. Since $(\widetilde{X}, \widetilde{Y})$ converges to $(\hat{X}, \hat{X})$ with $\hat{X} := (\hat{t}, \hat{x}, \hat{s}) \in O$, we may assume that for ϵ small enough, $(\widetilde{X}, \widetilde{Y})$ lies in O. We may write $\widetilde{X} := (t_1, x_1, s_1)$ and $\widetilde{Y} := (t_2, x_2, s_2)$. Then we have

$$\nabla_{\widetilde{X}}\phi_1 = -\nabla_{\widetilde{Y}}\phi_1 = \frac{2}{\epsilon}(\widetilde{X} - \widetilde{Y}).$$

Applying Crandall-Ishii's lemma (see Crandall et al. (1992)), we have that there exist Q and Q' in $\mathcal{S}^{N_{\bar{z}}+1}$ such that

$$\left(\nabla_{\widetilde{X}}\phi_1, Q\right) \in \bar{\mathcal{P}}^{1,2,\dots,2,+}\widetilde{W}_{\bar{z}}(\widetilde{X}), \quad \left(-\nabla_{\widetilde{Y}}\phi_1, Q'\right) \in \bar{\mathcal{P}}^{1,2,\dots,2,-}\widetilde{V}_{\bar{z}}^n(\widetilde{Y})$$

and the following matrix inequality holds in the non-negative definite sense:

$$\begin{pmatrix} Q & 0 \\ 0 & -Q' \end{pmatrix} \leq \frac{3}{\epsilon} \begin{pmatrix} I_{N_{\bar{z}}+1} & -I_{N_{\bar{z}}+1} \\ -I_{N_{\bar{z}}+1} & I_{N_{\bar{z}}+1} \end{pmatrix}.$$

By the viscosity subsolution (resp. supersolution) property of $\widetilde{W}$ (resp. $\widetilde{V}^n$), we have

$$\widetilde{F}_{\bar{z}}\left(t_1, x_1, s_1, \widetilde{W}, \nabla_{\widetilde{X}}\phi_1, Q\right) \leq 0 \tag{2.31}$$

and

$$\widetilde{F}_{\bar{z}}\left(t_2, x_2, s_2, \widetilde{V}^n, -\nabla_{\widetilde{Y}}\phi_1, Q'\right) \geq 0. \tag{2.32}$$

Subtracting (2.31) from (2.32), using the fact that the difference of the supreme is

less than the supreme of the difference, we obtain

$$\Gamma\left(\widetilde{W}_{\bar{z}}(\widetilde{X}) - \widetilde{V}_{\bar{z}}^n(\widetilde{Y})\right) + \sum_{i \in I_{\bar{z}}} \left(h_{\bar{z}}^i(s_1)\widetilde{W}_{\bar{z}}(\widetilde{X}) - h_{\bar{z}}^i(s_2)\widetilde{V}_{\bar{z}}^n(\widetilde{Y})\right)$$

$$\leq \sup_{\pi \in A} \left\{ J_1(\pi) + J_2(\pi) + J_3(\pi) \right\},$$

where

$$J_1(\pi) = (r + \theta^T \pi)\frac{2(x_1 - x_2)^2}{\epsilon} + \sum_{i \in I_{\bar{z}}} \mu_i \frac{2(s_{1i} - s_{2i})^2}{\epsilon},$$

$$J_2(\pi) = \sum_{i \in I_{\bar{z}}} \left(h_{\bar{z}}^i(s_1)\widetilde{W}_{\bar{z}i}\left(t_1, x_1\left(1 - \sum_{j=1}^N L_{ji}\pi^j\right), s_1^i\right)\right.$$

$$\left. - h_{\bar{z}}^i(s_2)\widetilde{V}_{\bar{z}i}^n\left(t_2, x_2\left(1 - \sum_{j=1}^N L_{ji}\pi^j\right), s_2^i\right)\right),$$

and

$$J_3(\pi) = \frac{1}{2}\pi^T\Sigma\pi\left(x_1^2 Q_{1,1} - x_2^2 Q'_{1,1}\right) + \frac{1}{2}\sum_{i \in I_{\bar{z}}}\sigma_i^2\left(s_{1i}^2 Q_{k_i,k_i} - s_{2i}^2 Q'_{k_i,k_i}\right)$$

$$+ \sum_{i,j \in I_{\bar{z}}, i<j}\rho_{ij}\sigma_i\sigma_j\left(s_{1i}s_{1j}Q_{k_i,k_j} - s_{2i}s_{2j}Q'_{k_i,k_j}\right)$$

$$+ \sum_{i \in I_{\bar{z}}}\rho_i^T\sigma\pi\sigma_i\left(x_1 s_{1i}Q_{1,k_i} - x_2 s_{2i}Q'_{1,k_i}\right).$$

Since $\phi_1(\widetilde{X}, \widetilde{Y}) \to 0$, we can derive $\limsup_{\epsilon \to 0} J_1(\pi) = 0$ for any π. By the definition of M, we have $\limsup_{\epsilon \to 0} J_2(\pi) \leq \sum_{i \in I_{\bar{z}}} h_{\bar{z}}^i(\hat{s})M$ for any π. By the structure condition and Crandall Ishii's inequality, we have

$$J_3(\pi) \leq \frac{K}{\epsilon}\left(|x_1 - x_2|^2 + \sum_{i \in I_{\bar{z}}}|s_{1i} - s_{2i}|^2\right).$$

Thus we can derive that $\limsup_{\epsilon \to 0} J_3(\pi) \leq 0$ for any π. Therefore

$$\limsup_{\epsilon \to 0} \left(\Gamma \left(\widetilde{W}_{\bar{z}}(\widetilde{X}) - \widetilde{V}_{\bar{z}}^n(\widetilde{Y}) \right) + \sum_{i \in I_{\bar{z}}} \left(h_{\bar{z}}^i(s_1)\widetilde{W}_{\bar{z}}(\widetilde{X}) - h_{\bar{z}}^i(s_2)\widetilde{V}_{\bar{z}}^n(\widetilde{Y}) \right) \right)$$
$$= \Gamma M + \sum_{i \in I_{\bar{z}}} h_{\bar{z}}^i(\hat{s})M \leq \sum_{i \in I_{\bar{z}}} h_{\bar{z}}^i(\hat{s})M.$$

Since $\Gamma > 0$, we have $M \leq 0$, which is a contradiction to the assumption that $M > 0$. We conclude that $M \leq 0$, which implies $W_z \leq V_z$ for $\forall z \in I$ on $[0, T) \times (0, \infty)^{N_z+1}$. $\qquad \square$

2.9 CONCLUSIONS

In this chapter we consider a utility maximization problem with looping contagion risk. We assume that the default intensity of one company depends on the stock prices of other companies and the default of one company induces immediate drops in the stock prices of the other surviving companies. In addition to the verification theorem, we prove the value function is the unique viscosity solution of the HJB equation system. We also compare and analyse the statistical distributions of terminal wealth of log utility based on two optimal strategies, one using the full information of intensity process, the other a proxy constant intensity process. Our numerical tests show that, statistically, using trading strategies based on stock price dependent intensities would achieve higher return on average, especially when the difference of the stock dependent intensity and the proxy constant intensity is big, but could also be more volatile in extreme scenarios. There remain many open questions in utility maximization with contagion risk, for example, the BSDE simulation method for power utility. We leave these and other questions to future research.

3

Dynamic Portfolio Optimization of Contingent Convertible Bond

3.1 Introduction

Contingent convertible bond, also known as CoCo bond, is a hybrid security issued by banks as debt instruments. The mechanism of CoCo bond is different from traditional convertible bond in that the conversion from debt to equity is contingent, which means CoCo bond is automatically converted into equity according to some pre-defined trigger event. The two main elements to define a CoCo bond are the trigger event and the conversion price. The trigger event of a CoCo bond can be based either on a mechanical rule or on supervisors' discretion or the mixture of both. Under the mechanical rule, the conversion is activated when a pre-specified capital ratio of the issuing bank is below a fixed threshold. The most used capital

ratio is the common equity tier 1 capital (CET1) ratio. Under the discretionary rule, or called 'point of non-viability' (PONV) rule, the conversion is activated based on the regulator's judgment when they think the conversion is necessary to prevent the issuing bank's insolvency. When the conversion is activated, the CoCo bond value is either written down or converted from debt into equity. The conversion price determines the rate at which the CoCo bond market value is converted into equity.

After 2008 financial crisis, Basel III introduces stronger regulation rules to financial institutions. Under Basel III, all banks are required to maintain at least 4.5% (improved from 2% under Basel II) of common equity tier 1 capital (CET1) of their risk-weighted assets (RWAs). As CoCo bond has the nature of automatically converted from debt into equity in financial distressed period, it is believed to improve the issuing bank's ability to absorb loss. Therefore, the issuance of CoCo bond by banks is supported by regulators, and the CoCos are counted into the RWAs of the issuing bank. For the issuing banks, CoCo bond is not only used to meet Basel III regulation requirement, but also cheaper than equity to maintain due to its debt nature. Since the first CoCo bond issued by Lloyds Banking Group in December 2009, there has been a rapid increase in the CoCo bond market size. Up to June 2018, the CoCo market size has reached €180bn. The European big banks such as Barclays, BNP Paribas, HSBC and Santander are the main contributors of CoCo market.

The CoCo bond is becoming extremely attractive to investors in the current low interest rate market. Compared with a corporate bond which normally pays coupon below 5%, Coco bond usually pays the fixed annual coupon between 5% to 10%. The high coupon rate reflects the fact that a CoCo bond is exposed to higher credit risk than the standard corporate bond. For a standard corporate bond, the main risk comes from the default, which depends on the issuing firm's credit risk level. However, the dominating risk of a CoCo bond is the conversion mechanism itself. As the design of conversion mechanism is to help the issuing banks to strengthen

their capital, the conversion risk is naturally higher than the default risk. Therefore, none of the CoCo bonds have credit rating higher than BBB in the current market. However, the issuers of CoCo bonds are not poorly rated corporates but highly rated banks, thus it is believed by many fund investors that the spread between default and conversion risk has been overestimated and the current high coupon is an arbitrage opportunity. The view is supported by the case of Banco Popular which immediately defaults before the CoCo conversion is triggered. In June 2017, Banco Popular becomes the first CoCo example which is wiped out of its value due to the resolution into Santander. The failure of Banco Popular gives an example where the conversion of CoCo bond is not necessarily strictly before the default occurs. Moreover, in the case of Banco Popular's failure, the spillover into the rest of the market is very little, which eliminates the concern from both investors and regulators that there exists strong contagion across CoCo bond market. Compared with equity, the CoCo bond market produced a record performance with returns of about 18% in 2017, which is more than European bank shares.

There has been extensive research in CoCo bond pricing and risk management. The CoCo bond pricing is addressed by two main modelling approaches – structural approach and reduced-form approach. The structural approach work includes Spiegeleer and Schoutens (2012) transforming the accounting trigger into a stock price trigger by introducing an implied stock price barrier level s^*. The implied s^* is unknown but can be calibrated from the historical market prices. The conversion time is defined to be the first passage time that the stock price drops below s^*. Based on the same approach, Corcuera et al. (2013) further employs a so-called smile conform model assuming the stock price follows a Levy process, incorporating fatter tails and jumps. The main drawback of both work is that the stock price at conversion is fixed under their assumption, which breaks the randomness of conversion price and consequently the loss given conversion of CoCo bond in reality. Brigo, et al. (2015) models the bank default by AT1P model and then converts the firm value into the accounting ratio by a linear regression model. Leung and Kwok (2015)

models the log CET1 ratio Y_t as an OU process and define two trigger types – one trigger event and Parisian trigger. One trigger event means that the first passage time that Y_t hits the fixed barrier B and Parisian trigger means that Y_t stays under the non-viable state $G > B$ cumulatively over a certain period of time. Based on the Parisian feature, they design numerical method for CoCo pricing. Jang, et al. (2018) models the issuing bank's stock price S_t by a geometric Brownian motion and its RWA-per-share value L by a single random variable. The conversion time is defined by the first passage time that the ratio of S_t/L is less than a fixed threshold value α_0. Under this setup, they obtain closed-form formula for the CoCo bond price with a CET1 ratio trigger. They further model the post-conversion risk by introducing a regulatory default time. The regulatory default time is defined by the first passage time (after conversion) that $S_t/L' < \alpha_1$, where L' is the new RWA-per-share random variable after dilution (due to conversion). They compare the difference between CoCo prices with/without post-conversion risk taking into account. The reduced-form approach models the conversion by a pure jump model. As it only requires the specification of the conversion intensity and the jump magnitude of the stock price at the conversion time, it is easier to calibrate than the equity structural approach. Cheridito and Xu (2015) assumes that the conversion intensity is deterministic. Chung and Kwok (2016) is the first trying to combine structural approach and reduced-form approach. It defines two conversion triggers – accounting trigger and regulatory trigger. The accounting trigger is modeled by the first passage time of the capital ratio hitting the trigger threshold. The regulatory trigger is defined as a pure jump process, while the jump intensity depends on both capital ratio and stock price. However, Chung and Kwok (2016) assumes that CoCo bond holder receives a predefined number of stock shares of the issuing bank. Firstly, this setting is unrealistic in the real CoCo bond market. Secondly, the assumption underestimates the conversion loss. As the market stock price at conversion is random, it is likely under such setting that the CoCo bond holder receives equity which is worth more than the CoCo bond market value when conversion occurs. Due to the technical

70

limitation, Chung and Kwok assume that the stock price drops a fixed fraction only when the conversion is caused by regulatory trigger, but not by accounting trigger.

To the best of our knowledge, there has been no existing research in the literature on dynamic portfolio optimization of CoCo bond. However, there exists quite a few work on defaulable bond optimization problem in the literature. Hou and Jin (2002) derives a closed-form solution for an investor who optimally allocates her wealth among a defaultable bond, a default-free stock and a risk-free bank account in finite time horizon. They assume the utility function is power utility and the default of bond is modelled by reduced-form approach where the default intensity follows an OU process. Bielecki and Jang (2006) extends the work of Hou and Jin (2002) by explicitly modelling the recovered amount in default using the conditional diversification assumption of Jarrow et al. (2005). All the parameters, however, are assumed to be constant. Bo et al. (2010) considers an infinite horizon portfolio optimization problem. The investor can optimally allocate her wealth into a defaultable perpetual bond, a default-free stock and a risk-free bank account. The utility function is log utility. They assume that both the intensity and the default premium process depend on a common Brownian factor. Unlike Hou and Jin (2002), where the dynamic of the defaultable bond is derived from the closed-form bond price formula given by Duffie and Singleton (1999), they heuristically postulate the dynamic without stochastic (Brownian motion) term and prove the dynamic is correct only if all the parameters are constant, see Remark 3.4.3. Capponi and Figueroa-Lopez (2011) considers a portfolio optimization problem in a regime-switching market, where the investor can dynamically allocate her wealth among a defaultable bond, a default-free stock, and a risk-free bank account. All the market coefficients depend on the market regimes. They derive the dynamic of defaultable bond under regime-switching model via the closed-form pricing formula in Duffie and Singleton (1999). They give a verification theorem which shows the solution of HJB equation under smoothness assumption is the value function.

All of the above papers assume that the defaultable bond is zero-coupon bond. However, one significant feature of CoCo bond is that it pays higher coupon than the standard defaultable bond. Therefore, investigating the impact of coupon is crucial for CoCo bond modelling. In this paper we analyse the optimal trading strategy of CoCo bond, where the investor can dynamically allocate her wealth into a CoCo bond that pays a continuous annual coupon with fixed rate c, and a risk-free bank account. We use the reduced-form approach for CoCo pricing and assume that the conversion intensity is a function of coupon rate and the issuing bank's stock price. The idea behind this assumption is that the conversion probability information is contained in both the coupon level c and the capital structure of the issuing bank. Moreover, as the conversion occurs only when the issuing bank's capital ratio is low, we make the further assumption that there is no conversion risk (the intensity value is zero) when stock price is above a fixed implied barrier level s^*. Under this setting, the drawback of Chung and Kwok (2016) such that the conversion risk is underestimated naturally disappears. Under such model setting, we compare CoCo bond investment performance with investing into the issuing bank's stock. To the best of our knowledge, we are the first in the literature tackling the dynamic portfolio optimization problem of CoCo bond. We extend Duffie and Singleton (1999) approach to derive the closed-form CoCo bond pricing formula with continuous coupon. Based on the risk-neutral pricing formula, we derive the dynamic of CoCo bond price under physical measure. We prove the pre-conversion value function is the unique viscosity solution of the corresponding HJB equation. We analyse the loss given conversion of CoCo bonds with different contractual features and propose the idea in our paper that the conversion intensity is a function of coupon rate and issuing bank's stock price, and becomes zero when stock price is above the implied barrier level s^*. This setting combines the flexibility of reduced-form approach and the special feature of CoCo bond such that the conversion risk is reflected by the coupon rate level and the capital structure. We compare and analyse the statistical distributions of terminal wealth of log utility and power utility based

72

on two investment assets, one with CoCo bond, the other with stock issued by the same bank. Our simulation results show that, the CoCo bond holders bear much more loss than equity holders when conversion occurs. However, investing into CoCo bond gets more profit (mean) while bearing less market risk (volatility) as long as conversion does not occur. We also analyse the sensitivity of terminal wealth distribution to different market scenarios.

The rest of the chapter is organized as follows. In Section 3.2, we introduce CoCo bond structure and analyse the loss given conversion of different contractual features. In Section 3.3, we introduce the market model. In Section 3.4, we derive the CoCo bond pricing formula under risk-neutral measure. In Section 3.5, we derive the dynamic of CoCo bond under physical measure by change of measure. In Section 3.6, we describe the dynamic portfolio optimization problem. Section 3.7 defines the value function and the corresponding HJB equation and Section 3.8 contains the main theorem (Theorem 3.8.5) which states that the value function is the unique viscosity solution of the HJB equation. In Section 3.9, we derive the closed-form optimal control for log utility and give the numerical method to solve power utility, and perform numerical tests and statistical distribution analysis. Section 3.10 concludes the chapter.

3.2 CoCo Bond Structure

A CoCo bond contract is characterized by its conversion trigger event and conversion price. Table 3.1 shows 5 live Coco contracts in the current market. Taking the contract with Bloomberg ISIN ID CH0244100266 as an example, the CoCo bond is issued by UBS AG and pays a fixed 5.125% coupon annually. The CoCo bond was issued close to the par value, with maturity date 15/05/2024. It will be written down if the CET1 ratio of UBS AG falls below 5%.

Issuing bank	UBS	HSBC	LLOYDS	CreditSuisse	Barclays
ISIN	CH0244100266	US404280AT69	US539439AG42	XS1076957700	US06738EAB11
Coupon rate	5.125%	6.375%	7.50%	6.25%	6.625%
Coupon frequency	Annually	Semi-annually	Quarterly	Semi-annually	Quarterly
Issue Price	99.905	100.0	100.0	100.0	100.0
Maturity	15/05/2024	Perpetual	Perpetual	Perpetual	Perpetual
Callable	No	30/03/2025	27/06/2024	18/12/2024	15/09/2019
Accounting trigger	CET1 < 5%	CET1 < 7%	CET1 < 7%	CET1 < 5.125%	CET1 < 7%
Regulatory trigger	No	No	No	Yes (by FINMA)	No
Conversion Price	Written-down	$4.03	$1.072	Written-down	£1.65

Table 3.1: CoCo bond contracts (Source: Bloomberg)

In the current CoCo market, most of the CoCo bonds are perpetual bonds which are callable after 5 or 10 years. Some of the CoCo bonds are immediately written down when the trigger events occur, and the rest are converted into equity shares of the issuing bank, e.g. US539439AG42 (LLOYDS). We denote the equity-converted CoCo bond by EC-CoCo and the written-down CoCo bond by WD-CoCo in the rest of paper.

For an EC-CoCo, the conversion price determines the number of shares the CoCo holders can receive. Denote C_P the conversion price and F the face value, then the number of converted shares C_r is equal to F/C_P in principle. Denote the conversion time by τ, then the equity share price S_τ is usually not equal to $S_{\tau-}$ as the conversion of CoCo bond has an immediate impact on the stock price. On one hand, the information of CoCo conversion activated is a damage to investor's confidence, which will have a negative impact on its stock price. On the other hand, a sudden injection of new shares into the market causes dilution effect. Both result in the stock price immediately dropping down. In this paper, we assume that the loss of stock price only comes from dilution effect, and the number of shares in the market before conversion is M and number of CoCo bonds in the market is M_C, then the total equity is $MS_{\tau-}$. Due to the dilution effect, the stock price after conversion becomes

$$S_\tau = \frac{MS_{\tau-}}{M + M_C C_r}.$$

Thus the loss fraction of stock price due to dilution effect is

$$L^S = \frac{S_{\tau-} - S_\tau}{S_{\tau-}} = \frac{M_C C_r}{M + M_C C_r}.$$

Denote $\alpha = F\frac{M_C}{M}$, then the loss fraction of stock price for a EC-CoCo is

$$L^S_{EC} = \frac{\alpha}{\alpha + C_P}.$$

Thus the EC-CoCo holder receives $C_r S_{\tau-}(1 - L^S_{EC})$ after conversion. The loss given

conversion ratio (LGC) of an EC-CoCo holder can be expressed by

$$LGC = 1 - \frac{C_r}{P_{\tau-}} S_{\tau-} (1 - L_{EC}^S),$$

where $P_{\tau-}$ is the CoCo market value right before the conversion. Note that the loss given conversion is linearly decreasing with C_r, thus C_r plays a crucial role in determining the LGC of EC-CoCo holder. Although C_r is equal to F/C_P in principle, there is a Conversion Shares Offer Consideration rule which allows the issuing bank to reduce the conversion ratio C_r on its discretion. The rule is designed to protect the share holders from significant dilution effect. Therefore, we approximate the actual conversion ratio C_r by $P_{\tau-}/C_P$ when estimating the LGC.* We will show that one benefit of such approximation is that the loss of CoCo bond holder is non-negative as long as the stock price at conversion is below some implied threshold, which makes the model more reasonable. With the approximation that $C_r = P_{\tau-}/C_P$, we have very tractable LGC formula

$$LGC = 1 - \frac{S_{\tau-}}{C_P} (1 - L_{EC}^S) = 1 - \frac{S_{\tau-}}{\alpha + C_P}.$$

For a written-down CoCo, there is no dilution effect thus the loss of stock price $L_{WD}^S = 0$. When the conversion trigger is activated, $1 - R$ fraction of face value is written down.† Therefore, R can be treated as the recovery rate and the written-down fraction $1 - R$ is the loss given conversion of the WD-CoCo holder.

The conversion price C_P normally takes the following representation:

$$C_P = \max\{\beta_1 S_{\tau-}, \beta_2 S_F\},$$

where $\beta_1, \beta_2 \in \{0, 1\}$ are constant and S_F is a pre-determined constant. Taking

---

*It is generally believed that the CoCo market price is lower than the face value right before conversion.

†The fraction R is determined on issuing bank's discretion at conversion time.

US539439AG42 (LLOYDS) for example, the conversion price is $1.072, thus $\beta_1 = 0, \beta_2 = 1, S_F = \$1.072$. The stock price of LLOYDS was $0.96 on the CoCo issuing date. The conversion fixed price is determined to be higher than the stock price at issuing date because a higher conversion price help to get rid of the significant dilution effect if conversion occurs. Note that most equity-converted CoCo bonds use the fixed conversion price.

Next we analyze the loss of stock price L_{EC}^S and the LGC of a EC-CoCo under different conversion price settings. Depending on the β values, we have the following scenarios:

- $\beta_1 = 0, \beta_2 = 1$. The conversion price $C_P = S_F$, which is a fixed price determined at CoCo inception. The downwards jump size of stock price is a constant equal to

$$L_{EC}^S = \frac{\alpha}{\alpha + S_F},$$

and the loss given conversion of CoCo investor becomes a linear decreasing function of stock price before conversion

$$LGC = 1 - \frac{S_{\tau-}}{\alpha + S_F}.$$

Note that LGC in this case can be negative. A sufficient condition to guarantee that LGC is non-negative is $S_{\tau-} \leq \alpha + S_F$.

- $\beta_1 = 1, \beta_2 = 0$. The conversion price $C_P = S_{\tau-}$, which is the stock price at the conversion time. The loss fraction of stock price is a decreasing function of stock price before conversion

$$L_{EC}^S = \frac{\alpha}{\alpha + S_{\tau-}},$$

and the loss given conversion of CoCo investor becomes

$$LGC = 1 - \frac{S_{\tau-}}{\alpha + S_{\tau-}} = \frac{\alpha}{\alpha + S_{\tau-}}.$$

- $\beta_1 = 1, \beta_2 = 1$. The conversion price $C_P = \max\{S_{\tau-}, S_F\}$, which is the stock price at the conversion time floored by a constant S_F. The loss fraction of stock price is

$$L_{EC}^S = \frac{\alpha}{\alpha + \max\{S_{\tau-}, S_F\}},$$

and the loss given conversion becomes

$$LGC = 1 - \frac{S_{\tau-}}{\alpha + \max\{S_{\tau-}, S_F\}} = \frac{\alpha + \max\{S_{\tau-}, S_F\} - S_{\tau-}}{\alpha + \max\{S_{\tau-}, S_F\}}.$$

Remark 3.2.1. Note that under the contracts $\{\beta_1 = 1, \beta_2 = 0\}$ and $\{\beta_1 = 1, \beta_2 = 1\}$, it is naturally true that $LGC \in (0, 1]$. However, under the first case $\{\beta_1 = 0, \beta_2 = 0\}$, the loss given conversion could be negative if $S_{\tau-} > \alpha + S_F$. In this case, the CoCo holder gains money instead of losing money when CoCo bond is forced to convert. The model is not realistic if the stock price is still at high level when the conversion is close to be activated. In our market model, we force $S_{\tau-} \leq \alpha + S_F$ by introducing an implied stock price barrier s^* above which the CoCo bond is not exposed to conversion risk.

3.3 Market Model Setting

Assume that there are one contingent convertible (CoCo) bond whose maturity is T_1[‡] and one risk-free bank account in the market. We denote the CoCo bond price by $\{P_t\}_{t \geq 0}$, the risk-free bank account amount by $(B_t)_{t \geq 0}$ with the risk-free interest rate r. We assume that the CoCo bond is automatically written down or converted into

---

[‡]As most of the CoCo bonds in the market are perpetual bonds which are callable after 10 years, we assume T_1 is much greater than the portfolio optimization maturity T throughout this paper.

$C_r \in [0, \infty)$ shares of equity, whose price is denoted by $\{\tilde{S}_t\}_{t \geq 0}$, after a contractually pre-defined trigger event occurs, with loss given conversion (LGC) rate $L_t \in [0, 1]$. Before the conversion occurs, CoCo bond pays a fixed continuous coupon at constant annual rate c. Assume $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t \geq 0}, \mathbb{P})$ is a complete probability space satisfying the usual conditions. $(\mathcal{G}_t)_{t \geq 0}$ is an enlarged filtration given by $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$, where $\mathcal{F}_t$ is the filtration generated by a standard Brownian motion W and be independent of $(H_t)_{t \geq 0}$. Let τ be a non-negative random variable representing the conversion time of the CoCo bond, defined by

$$\tau := \inf \left\{ t \geq 0 : \int_0^t h_u du \geq \mathcal{X} \right\},$$

where $(h_t)_{t \geq 0}$ is the Hazard (intensity) process and $\mathcal{X}$ is a standard exponential random variable on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$ and is independent of the filtration $(\mathcal{F}_t)_{t \geq 0}$, which means τ are totally inaccessible stopping times. Then $(\mathcal{H}_t)_{t \geq 0}$ is defined by $\mathcal{H}_t := \sigma(H_u : u \leq t)$, where $H_t := \mathbb{I}_{\{\tau \leq t\}}$ which equals 0 if $\tau > t$ and 1 otherwise. The default indicator process $(H_t)_{t \geq 0}$ is associated with the intensity process $(h_t)_{t \geq 0}$. The consequence of such construction of τ is the so-called $H-$ hypothesis, which means any $\mathcal{F}$–square integrable martingale is also a $\mathcal{G}$-square integrable martingale, see Bielecki and Rutkowski (2003). The market model before conversion occurs is given by the following SDEs (stochastic differential equations):

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t, \quad \frac{dB_t}{B_t} = rdt, \tag{3.1}$$

where μ and σ are drift and volatility of the issuing bank's stock price. When conversion occurs, we assume that the stock price drops immediately by a fraction $L^S \in [0, 1)$. Therefore, the traded stock price $\tilde{S}_t$ has the following dynamic:

$$\frac{d\tilde{S}_t}{\tilde{S}_{t-}} = \mu dt + \sigma dW_t - L_t^S dH_t.$$

Based on the analyses in Section 3.2, we make the following assumptions.

79

Assumption 3.3.1. The loss of stock price at conversion $L_t^S := L^S(S_t)$ where $L^S(\cdot)$ is a bounded deterministic function.

Assumption 3.3.2. The loss given conversion of CoCo bond $L_t := L(S_t)$ is a bounded, continuous deterministic function.

Remark 3.3.3. Note that S_t and $\tilde{S}_t$ are identical up to conversion time τ. Due to technical reason, we assume L_t^S and L_t are functions of S_t instead of $\tilde{S}_{t-}$, as L_t^S and L_t impact the optimization problem only up to (including) conversion. After conversion, the CoCo bond value drops to zero and stock price follows the standard Geometric Brownian motion. Under the current assumption, both L_t^S and L_t are $\mathcal{F}$–measurable.

3.4 CoCo Bond Pricing under Risk-neutral Measure

We derive the CoCo bond pricing formula based on the model setting specified in Section 3.3. For simplicity, we assume the face value $F = 1$ throughout this paper. Note that under Assumption 3.3.2, L_t is $\mathcal{F}$–predictable. We treat CoCo bond as a bond component while the equity component is only a recovery amount. After conversion, the CoCo bond has zero value. Duffie and Singleton (1999) derives the pricing formula of a zero-coupon defaultable bond P_t^0 under recovery of market value (RMV) scheme:

$$P_t^0 = \mathbb{I}_{\{\tau > t\}} \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^{T_1} (r + L_u h_u^{\mathbb{Q}}) du} \mid \mathcal{F}_t \right],$$

where the $\mathcal{F}$–measurable process $h^{\mathbb{Q}}$ is the conversion intensity under risk-neutral measure $\mathbb{Q}$. We will discuss the relation between $h^{\mathbb{Q}}$ and h in the following sections. For technical reason, we assume that the continuous coupon is paid based on the CoCo market value instead of the face value F. Following the virtue of Duffie and Singleton (1999), under recovery of market value (RMV) scheme, we heuristically

conjecture the CoCo pricing formula as following:

$$P_t = \mathbb{I}_{\{\tau > t\}} \mathbb{E}^{\mathbb{Q}} \left[e^{- \int_t^{T_1} (r - c + L_u h_u^{\mathbb{Q}}) du} \mid \mathcal{F}_t \right]. \tag{3.2}$$

Following the similar proof of Duffie and Singleton, in order to confirm this conjecture, we use the fact that the discounted gain process must be a martingale under $\mathbb{Q}$. The discounted gain process G is defined by

$$G_t := e^{-rt} V_t (1 - H_t) + \int_0^t e^{-rs} (1 - L_s) V_{s-} dH_s + \int_0^t e^{-rs} c V_s (1 - H_s) ds,$$

where V_t is the pre-default CoCo price. As Duffie and Singleton, we suppose that V does not itself jump at the default time τ, then applying Ito's Formula to G_t gives

$$dG_t = e^{-rt} (1 - H_t) \left(dV_t - (r - c + L_t h_t^{\mathbb{Q}}) V_t dt - L_t V_t (dH_t - h_t^{\mathbb{Q}} (1 - H_t) dt) \right).$$

Thus for G to be a $\mathbb{Q}$ martingale, it is necessary and sufficient that

$$V_t = \int_0^t (r - c + L_t h_t^{\mathbb{Q}}) V_t dt + M_t$$

for some $\mathbb{Q}$-martingale M_t. Given the terminal condition $V_{T_1} = 1$, this implies (3.2) by Duffie and Singleton (1999).

Based on the pricing formula (3.2), we derive the dynamic of CoCo bond price under risk-neutral measure $\mathbb{Q}$ by the following proposition.

Proposition 3.4.1. *Assuming that the intensity process $h_t^{\mathbb{Q}}$ is $\mathcal{F}$–measurable and satisfies $\mathbb{E} \left[e^{-2 \int_0^{T_1} (r + L_u h_u^{\mathbb{Q}}) du} \right] < \infty$, then the dynamic of CoCo bond price under $\mathbb{Q}$ is given by*

$$\frac{dP_t}{P_{t-}} = \left(r - c + L_t h_t^{\mathbb{Q}} \right) dt + \lambda_t dW_t^{\mathbb{Q}} - dH_t, \tag{3.3}$$

where λ_t is a $\mathcal{F}$–predictable process.

Proof. Denote $V_t := \mathbb{E}^{\mathbb{Q}} \left[e^{- \int_t^{T_1} (r - c + L_u h_u^{\mathbb{Q}}) du} \mid \mathcal{F}_t \right]$, then (3.2) can be written as $P_t = $

$\mathbb{I}_{\{\tau > t\}} V_t$. We reformat V_t into $V_t = b_t \phi_t$, where $b_t := e^{\int_0^t (r - c + L_u h_u^{\mathbb{Q}}) du}$ and

$$\phi_t := \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_0^{T_1} (r - c + L_u h_u^{\mathbb{Q}}) du} \mid \mathcal{F}_t \right]. \tag{3.4}$$

Since ϕ_t is a $(\mathbb{Q}, \mathcal{F})$–martingale by conditional expectation structure, and the random variable $e^{-\int_0^{T_1} (r + L_u h_u^{\mathbb{Q}}) du}$ is square-integrable and $\mathcal{F}$–measurable, there exists a $\mathcal{F}$–predictable process λ_t such that

$$d\phi_t = \phi_t \lambda_t dW_t^{\mathbb{Q}},$$

by Martingale representation theorem. Thus the pre-conversion CoCo price V_t follows dynamic

$$\frac{dV_t}{V_t} = (r - c + L_t h_t^{\mathbb{Q}}) dt + \lambda_t dW_t^{\mathbb{Q}}.$$

The dynamic of CoCo bond price is given by

$$
\begin{aligned}
dP_t &= (1 - H_{t-}) dV_t - V_{t-} dH_t \\
&= P_{t-} \left((r - c + L_t h_t^{\mathbb{Q}}) dt + \lambda_t dW_t^{\mathbb{Q}} \right) - P_{t-} dH_t,
\end{aligned}
$$

which implies

$$\frac{dP_t}{P_{t-}} = \left(r - c + L_t h_t^{\mathbb{Q}} \right) dt + \lambda_t dW_t^{\mathbb{Q}} - dH_t.$$

$\square$

Remark 3.4.2. By Martingale representation theorem, the volatility of CoCo bond dynamic λ_t is an unknown $\mathcal{F}$–predictable process. We assume λ a deterministic function of t, c later in this chapter. In the numerical test section, we will treat it as a constant volatility and calibrate to the historical CoCo prices.

Remark 3.4.3. Bo et al. (2010) conjecture the price dynamics for a defaultable perpetual bond that pays a constant coupon $\widetilde{C}$ per unit time as follows:

$$dP_t = r P_t dt + \frac{L h_t}{\eta_t} P_t dt - \widetilde{C} dt - L dH_t, \tag{3.5}$$

where $\frac{1}{\eta_t}$ is the default risk premium. They verify that the proposed dynamic is correct if all the parameters are constant. For their market model where default intensity and risk premium are both deterministic functions of an external factor, however, they directly randomize the market parameters of (3.5) and mention that the dynamic might be invalid, see Remark 2.1 in Bo et al. (2010). If all the parameters are constant in our model setting, we have ϕ_t defined by (3.4) constant, thus $\lambda_t = 0$ as a special case of Martingale representation theorem. Thus the CoCo dynamic (3.3) degenerates to (3.5). However, note that directly randomizing the market parameters of (3.5) under heuristic argument loses the stochastic term $\lambda dW_t^{\mathbb{Q}}$.

3.5 Dynamic of CoCo Bond under Physical Measure

We have derived the dynamic of CoCo bond under risk-neutral measure. However, the dynamic portfolio optimization problem is defined under physical measure $\mathbb{P}$. We firstly discuss change of measure in this section, and derive the dynamic of CoCo bond under physical measure accordingly. All the following discussions are based on Bielecki and Rutkowski (2003).

Let $\mathbb{Q}$ be an equivalent martingale measure. We work under the H–hypothesis. Let us fix the investment terminal time by $T > 0$ and $T \ll T_1$. We introduce the Radon-Nikodym density process η_t for any $t \in [0, T]$ by

$$\eta_t := \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{G}_t} = \mathbb{E}[\eta_T \mid \mathcal{G}_t],$$

where η_T is a $\mathcal{G}_T$–measurable and integrable random variable, such that $\mathbb{P}(\eta_T > 0) = 1$ and $\mathbb{E}[\eta_T] = 1$. η_t is a $(\mathbb{P}, \mathcal{G})$–martingale by construction. Thus according to Bielecki and Rutkowski (2003) (Corollary 5.2.4), η admits the following dynamic

representation

$$d\eta_t = \eta_{t-}\left(\beta_t dW_t + \kappa_t dM_t\right),$$

where β_t and κ_t are $\mathcal{G}$–predictable processes, and M_t is a $(\mathbb{P}, \mathcal{G})$–martingale given by

$$M_t = H_t - \int_0^t (1 - H_{u-})h_u du.$$

Further, by Bielecki and Rutkowski (2003) (Proposition 5.3.1), the process

$$W_t^{\mathbb{Q}} = W_t - \int_0^t \beta_u du$$

follows a Brownian motion with respect to $\mathcal{G}$ under $\mathbb{Q}$, and the process

$$M_t^{\mathbb{Q}} = H_t - \int_0^t (1 - H_{u-})(1 + \kappa_u)h_u du$$

follows a $\mathcal{G}$–martingale under $\mathbb{Q}$. Therefore, the relation between $h^{\mathbb{Q}}$ and h is given by $h_t^{\mathbb{Q}} := (1 + \kappa_t)h_t$. The quantity $1 + \kappa_t$ is the coverage ratio, which reflects the conversion risk premium. Empirically, κ_t decreases with conversion risk and converges to zero when conversion risk tends to infinity, see Heynderickx et al. (2016). The dynamic of traded stock price under $\mathbb{Q}$ is

$$\frac{d\tilde{S}_t}{\tilde{S}_{t-}} = \left(\mu + \sigma\beta_t - L_t^S(1 + \kappa_t)h_t\right) dt + \sigma dW_t^{\mathbb{Q}} - L_t^S M_t^{\mathbb{Q}}.$$

Since the drift of stock price is equal to interest rate r under risk-neutral measure $\mathbb{Q}$, we have the following relation among processes β_t, κ_t and h_t:

$$\beta_t = \frac{r - \mu + L_t^S(1 + \kappa_t)h_t}{\sigma}.$$

Combining with Proposition 3.4.1, we get the dynamic of CoCo bond under physical measure $\mathbb{P}$ by the following proposition.

Proposition 3.5.1. *Assuming that both h_t and κ_t are $\mathcal{F}$–measurable and satisfy* $\mathbb{E}\left[e^{-2\int_0^{T_1}(r+L_u(1+\kappa_u)h_u)du} \right] < \infty$, *then the dynamic of CoCo bond price under $\mathbb{P}$ is given by*

$$\frac{dP_t}{P_{t-}} = \left(r - c + L_t(1+\kappa_t)h_t - \beta_t\lambda_t \right) dt + \lambda_t dW_t - dH_t.$$

Proof. As both h_t and κ_t are $\mathcal{F}$–measurable and satisfy $\mathbb{E}\left[e^{-2\int_0^{T_1}(r+L_u(1+\kappa_u)h_u)du} \right] < \infty$, the conditions in Proposition 3.4.1 are satisfied and

$$\frac{dP_t}{P_{t-}} = \left(r - c + L_t(1+\kappa_t)h_t \right) dt + \lambda_t dW_t^{\mathbb{Q}} - dH_t.$$

Replacing $dW_t^{\mathbb{Q}}$ by $dW_t - \beta_t dt$, we get the desired dynamic under $\mathbb{P}$. $\qquad\square$

Remark 3.5.2. Define $\theta_t := L_t(1+\kappa_t)h_t - \beta_t\lambda_t$, then the drift of CoCo bond under physical measure $\mathbb{P}$ is $r - c + \theta_t$. Since we assume that CoCo bond pays annual coupon at rate c, the yield of CoCo bond is then $r + \theta_t$. Thus θ_t is the exceed return over risk-free rate r. As the return of current CoCo market is greater than that of the equity market, we have $r + \theta_t > \mu$. We will use this relation in the numerical test section.

3.6 DESCRIPTION OF PORTFOLIO OPTIMIZATION PROBLEM

Investors dynamically allocate proportion $(\pi, 1 - \pi)$ of their total wealth into CoCo bond P and bank account B before conversion occurs. After conversion, CoCo bond is converted into equities, thus investors dynamically allocate proportion $(\pi, 1 - \pi)$ of their total wealth into stock S and bank account B.

The admissible control set $\mathcal{A}$ is the set of control processes π that are progressively measurable with respect to the filtration $(\mathcal{F}_t)$ and $\pi_t \in A$ for all $t \in [0, T]$. The set

A is defined by

$$A := \big\{\pi \in O \text{ and } \pi \leq 1 - \epsilon_A\big\}$$

for some bounded set $O \subset \mathbb{R}$ and constant $\epsilon_A > 0$. Note that the condition $\pi \leq 1 - \epsilon_A$ guarantees that $1 - L_t \pi \geq \epsilon_A$ for $t \in [0, T]$. Therefore the wealth process is always non-negative.

We can naturally split the portfolio optimization problem into two stages – pre-conversion stage and post-conversion stage. Denote the wealth process by $\{X_t\}_{t \geq 0}$, then the dynamic of pre-conversion wealth process is given by

$$
\begin{aligned}
\frac{dX_t}{X_{t-}} &= \pi_{t-}\left(\frac{dP_t}{P_{t-}} + c\,dt + (1 - L_t)dH_t\right) + (1 - \pi_t)\frac{dB_t}{B_t} \\
&= (r + \theta_t \pi_t)\,dt + \lambda_t \pi_t dW_t - L_t \pi_{t-} dH_t,
\end{aligned}
\tag{3.6}
$$

and the dynamic of post-conversion wealth process is given by

$$
\frac{dX_t}{X_t} = (r + (\mu - r)\pi_t)\,dt + \sigma \pi_t dW_t,
$$

which is a standard Merton's problem with default-free stock and bank account.

Remark 3.6.1. For a given control process $\pi \in \mathcal{A}$, equation (3.6) admits a unique strong solution that satisfies

$$
\sup_{t \in [0,T]} \mathbb{E}\left[X_t^\alpha\right] \leq K x^\alpha
\tag{3.7}
$$

for any $\alpha > 0$. This can be easily verified as $X_t^\alpha = x^\alpha N_t M_t$, where

$$
\begin{aligned}
N_t &:= \exp\left(\alpha \int_0^t \left(r + \theta_u \pi_u - \frac{1}{2}\lambda_u^2 \pi_u^2\right) du + \frac{1}{2}\alpha^2 \int_0^t \lambda_u^2 \pi_u^2 du \right. \\
&\qquad\quad \left. + \alpha \int_0^t \ln(1 - L_u \pi_{u-})dH_u\right), \\
M_t &:= \exp\left(\int_0^t \alpha \lambda_u \pi_u dW_u - \frac{1}{2}\alpha^2 \int_0^t \lambda_u^2 \pi_u^2 du\right).
\end{aligned}
$$

Since A_t is a bounded set for any $t \in [0, T]$ and $1 - L_t \pi \geq \epsilon_A$, we have $|N_t| < K$, independent of t, and M_t is an exponential martingale, thus $\mathbb{E}[M_t] = 1$, which gives (3.7).

3.7 VALUE FUNCTION AND HAMILTON-JACOBI-BELLMAN EQUATION

Our objective is to maximize the expected utility of the terminal wealth, that is,

$$\sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U\left(X_T^\pi\right)\right].$$

We make the following assumption of the general utility function. Note that the Power utility $U(x) = (1/\gamma)x^\gamma$, $0 < \gamma < 1$, satisfies Assumption 3.7.1, but not log utility. However, due to the special structure of log utility, we derive the optimality and regularity of the value function directly in the following numerical test.

Assumption 3.7.1. The utility function U is defined on $[0, \infty)$, is continuous, non-decreasing, concave, and satisfies $U(0) > -\infty$ and $|U(x)| \leq K(1 + x^\gamma)$ for all $x \in [0, \infty)$, where $K > 0$ and $0 < \gamma < 1$ are constants.

In the above sections, we require the processes h_t, κ_t to be $\mathcal{F}$–measurable and satisfy some square integrable conditions. To make the optimization problem adapt to the proper HJB equation, we make the further assumptions:

Assumption 3.7.2. We make the following assumptions on the processes $h_t, \kappa_t, \lambda_t, \theta_t$:

- $h_t := h(c, S_t)$ is a bounded deterministic function that is continuous in S_t for $\forall c > 0$. Further, $h(c, s) = 0$ if $s \geq s^*$ where $s^* \in (0, S_F + \alpha]$ is an implied stock price barrier.

- $\kappa_t := \kappa(t, c)$ is a bounded deterministic function that is continuous in $t \in [0, T)$ for $\forall c > 0$.

87

- $\lambda_t := \lambda(t, c)$ is a bounded deterministic function that is Lipschitz continuous in $t \in [0, T)$ for $\forall c > 0$.

- $\theta_t := \theta(t, c, S_t)$ is Lipschitz continuous in $t, s \in [0, T) \times (0, \infty)$ for $\forall c > 0$.

Remark 3.7.3. We assume that the coupon rate c is fixed throughout this paper, thus the constant c can be seen as a correction term of the functions defined in Assumption 3.7.2. Intuitively, both h and θ are increasing with c. Empirically, κ is decreasing with c, see Heynderickx et al. (2016). Examples can be found in numerical tests section.

Remark 3.7.4. Under the assumption that $h(c, s) = 0$ if $s \geq s^*$, it is practically reasonable that the CoCo bond is not exposed to conversion risk when the stock price is still at a high level. The parameter s^* is a fixed implied stock price barrier level, which can be calibrated to the historical data, see Spiegeleer and Schoutens (2012).

Due to the existence of conversion event, the problem can be naturally split into pre-conversion case and post-conversion case. The latter is a standard Merton's problem and the post-conversion value function v_1 is a function of time t and wealth x only. We will only discuss the pre-conversion case in the following sections. The pre-conversion value function v_0 is defined by

$$v_0(t, x, s) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U\left(X_T^\pi\right) \mid X_t = x, S_t = s, H_t = 0\right]$$

for $(t, x, s) \in [0, T] \times (0, \infty)^2$. Note that the coupon rate c is assumed to be constant, thus not one of the variables of v_0. Moreover, if we assume h is independent of s, then v_0 is a function of t, x only.

Remark 3.7.5. Combining Assumption 3.7.1 and Remark 3.6.1, we have

$$|v_0(t, x, s)| \leq K(1 + x^\gamma).$$

The pre-conversion value function satisfies the following HJB equation by the dynamic programming principle

$$-\sup_{\pi \in A} \mathcal{L} v_0(t, x, s) = 0 \tag{3.8}$$

for $(t, x, s) \in [0, T) \times (0, \infty)^2$ with terminal condition $v_0(T, x, s) = U(x)$, where $\mathcal{L}$ is the infinitesimal generator of S and X processes with control π, given by

$$
\begin{aligned}
\mathcal{L} v_0(t, x, s) \;=\;& \frac{\partial v_0}{\partial t} + rx \frac{\partial v_0}{\partial x} + \mu s \frac{\partial v_0}{\partial s} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 v_0}{\partial s^2} - h(c, s) v_0 \\
& + \theta(t, c, s)\pi x \frac{\partial v_0}{\partial x} + \frac{1}{2}\lambda(t, c)^2 \pi^2 x^2 \frac{\partial^2 v_0}{\partial x^2} \\
& + \sigma \lambda(t, c)\pi x s \frac{\partial^2 v_0}{\partial x \partial s} + h(c, s) v_1 \left(t, x(1 - L(s)\pi)\right), \tag{3.9}
\end{aligned}
$$

where v_1 is the known post-conversion value function.

3.8 Viscosity Solution Representation

As post-conversion case is a classical Merton's problem, we assume that the post-conversion value function $v_1 \in C^{1,2}\left([0, T) \times (0, \infty)\right) \cup C\left([0, T] \times (0, \infty)\right)$. To facilitate discussions of viscosity solution, we define F function by

$$F\left(t, x, s, w, \nabla_{(t,x,s)} w, \nabla^2_{(x,s)} w\right) = -\sup_{\pi \in A} \mathcal{L} w(t, x, s),$$

where $\nabla_{(t,x,s)} w \in \mathbb{R}^3$ is the gradient vector of w with respect to (t, x, s), and $\nabla^2_{(x,s)} w \in \mathbb{R}^{2 \times 2}$ is the Hessian matrix of w with respect to (x, s). w and its derivatives are evaluated at (t, x, s). The HJB equation (3.8) is the same as

$$F\left(t, x, s, v_0, \nabla_{(t,x,s)} v_0, \nabla^2_{(x,s)} v_0\right) = 0.$$

We denote that w^* and w_* are upper-semicontinuous envelope and lower-semicontinuous envelope of w respectively. We will show in this section that the pre-conversion value function v_0 is the unique viscosity solution to equation (3.8) based on the definition below.

Definition 3.8.1. (i) w is a (discontinuous) viscosity subsolution of (3.8) on $[0, T) \times (0, \infty)^2$ if

$$F\left(\bar{t}, \bar{x}, \bar{s}, \varphi, \nabla_{(t,x,s)}\varphi, \nabla^2_{(x,s)}\varphi\right) \leq 0$$

for all $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty)^2$ and $\varphi \in C^{1,2,2}\left([0, T) \times (0, \infty)^2\right)$ such that $w^*(\bar{t}, \bar{x}, \bar{s}) = \varphi(\bar{t}, \bar{x}, \bar{s})$ and $w^* \leq \varphi$ on $[0, T) \times (0, \infty)^2$.

(ii) w is a (discontinuous) viscosity supersolution of (3.8) on $[0, T) \times (0, \infty)^2$ if

$$F\left(\bar{t}, \bar{x}, \bar{s}, \varphi, \nabla_{(t,x,s)}\varphi, \nabla^2_{(x,s)}\varphi\right) \geq 0$$

for all $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty)^2$ and $\varphi \in C^{1,2,2}\left([0, T) \times (0, \infty)^2\right)$ such that $w_*(\bar{t}, \bar{x}, \bar{s}) = \varphi(\bar{t}, \bar{x}, \bar{s})$ and $w_* \geq \varphi$ on $[0, T) \times (0, \infty)^2$.

(iii) We say that w is a (discontinuous) viscosity solution of (3.8) on $[0, T) \times (0, \infty)^2$ if it is both a viscosity subsolution and supersolution of (3.8).

Note that in Chapter 1, the viscosity solution is defined by the solution of the HJB equation system. In this chapter, we have that v_1 is smooth in prior, thus the viscosity solution definition 3.8.1 can be treated as a special (simplified) version of the definition 2.6.1 in Chapter 1.

The main result in this section is that the pre-conversion value function v_0 is the unique viscosity solution of equation (3.8) on $[0, T) \times (0, \infty)^2$ if v_0 is continuous on the boundary. This result is given by the following two propositions.

Proposition 3.8.2. *The pre-conversion value function v_0 is a viscosity solution of equation (3.8) on $[0, T) \times (0, \infty)^2$.*

The uniqueness of viscosity solution is given by the following comparison principle result.

Proposition 3.8.3. *Let W (resp. V) be a u.s.c. viscosity subsolution (resp. l.s.c. viscosity supersolution) of (3.8) on $[0,T) \times (0,\infty)^2$ and satisfy the growth condition $|W|, |V| \leq K(1 + x^\gamma)$ for $\gamma > 0$, the terminal relation $W(T,x,s) \leq V(T,x,s)$, and the boundary relations $W(t,0,s) \leq V(t,0,s)$ and $W(t,x,0) \leq V(t,x,0)$ for $\forall (t,x,s) \in [0,T) \times [0,\infty)^2$. Then we have $W \leq V$ on $[0,T] \times [0,\infty)^2$.*

We skip the proofs of Proposition 3.8.2 and 3.8.3 in this chapter, as they are special cases of the viscosity results in Chapter 1.

Remark 3.8.4. The boundedness and Lipschitz continuity assumptions in Assumption 3.7.2 are essential in comparison principle proof. Moreover, we assume the structure condition in the comparison principle proposition in Chapter 1. We slightly modify the definition of structure condition in this chapter by

$$J(\pi) \leq \frac{K}{\epsilon} \left(|t_1 - t_2|^2 + |x_1 - x_2|^2 + |s_1 - s_2|^2 \right), \ \forall \pi \in A,$$

where

$$J(\pi) = \frac{1}{2} v^T \begin{pmatrix} Q & 0 \\ 0 & -Q' \end{pmatrix} v,$$

where

$$v = \left(\lambda(t_1, c)\pi x_1, \sigma s_1, \lambda(t_2, c)\pi x_2, \sigma s_2 \right)^T,$$

for $(t_i, x_i, s_i) \in \mathcal{O}$ such that $\mathcal{O} \in [0,T) \times (0,\infty)^2$ is bounded for $i = 1, 2$ and Q, Q' satisfy

$$\begin{pmatrix} Q & 0 \\ 0 & -Q' \end{pmatrix} \leq \frac{3}{\epsilon} \begin{pmatrix} I_2 & -I_2 \\ -I_2 & I_2 \end{pmatrix}.$$

Actually, simple algebraic calculation shows that

$$J(\pi) \leq \frac{3}{2\epsilon} \upsilon^T \begin{pmatrix} I_2 & -I_2 \\ -I_2 & I_2 \end{pmatrix} \upsilon$$

$$\leq \frac{3}{2\epsilon} \left(\pi^2 |\lambda(t_1, c)x_1 - \lambda(t_2, c)x_2|^2 + \sigma^2 |s_1 - s_2|^2 \right).$$

By Lipschitz continuity of $\lambda(t, c)$, we have

$$J(\pi) \leq \frac{3}{2\epsilon} \left(\pi^2 \lambda(t_1, c)^2 |x_1 - x_2|^2 + \pi^2 x_2^2 |\lambda(t_1, c) - \lambda(t_2, c)|^2 + \sigma^2 |s_1 - s_2|^2 \right)$$

$$\leq \frac{3}{2\epsilon} \left(\pi^2 \lambda(t_1, c)^2 |x_1 - x_2|^2 + \pi^2 x_2^2 K |t_1 - t_2|^2 + \sigma^2 |s_1 - s_2|^2 \right).$$

Due to boundedness of $\lambda(t, c)$ and x_i, the structure condition is satisfied.

We finally conclude the main theorem of this section:

Theorem 3.8.5. *Suppose that the pre-conversion value function v_0 satisfies the terminal condition $(v_0)^*(T, x, s) \leq (v_0)_*(T, x, s)$ and boundary conditions $(v_0)^*(t, 0, s) \leq (v_0)_*(t, 0, s), (v_0)^*(t, x, 0) \leq (v_0)_*(t, x, 0)$ for $\forall(t, x, s) \in [0, T] \times [0, \infty)^2$, then it is the unique viscosity solution of equation (3.8) on $[0, T) \times (0, \infty)^2$.*

3.9   NUMERICAL TESTS

In this section, we give an example with log utility and perform some numerical tests and statistical and sensitivity analysis.

3.9.1 OPTIMAL STRATEGIES FOR LOG UTILITY

For $U(x) = \ln x$, the post-conversion case is well known with the optimal control $\pi = (\mu - r)/(\sigma)^2$ and the post-conversion value function $v_1(t, x) = \ln x + g_1(t)$, where

$$g_1(t) = \left(r + \frac{1}{2} \left(\frac{\mu - r}{\sigma} \right)^2 \right) (T - t).$$

We conjecture that the pre-conversion value function takes the form

$$v_0(t, x, s) = \ln x + f(t, s). \tag{3.10}$$

Substituting (3.10) into (3.8), we get a linear PDE for f:

$$\frac{\partial f}{\partial t} + \frac{1}{2} (\sigma s)^2 f_{ss} + \mu s f_s - h(c, s) f + r + h(s) g_1(t) + \sup_{\pi \in A} G(t, c, s, \pi) = 0 \tag{3.11}$$

with the terminal condition $f(T, s) = 0$, where G is defined by

$$G(t, c, s, \pi) = \theta(t, c, s)\pi - \frac{1}{2}\lambda(t, c)^2 \pi^2 + h(c, s) \ln \left(1 - L(s)\pi \right).$$

Since A is compact and G is continuous, there exists an optimal solution which satisfies the optimality condition

$$\theta(t, c, s) - \lambda(t, c)^2 \pi - L(s)h(c, s)\frac{1}{1 - L(s)\pi} = 0 \tag{3.12}$$

and we get the optimal control π^* as

$$(\pi)^*(t, c, s) = \min \left\{ (\pi)_f^*(t, c, s), \frac{1 - \epsilon_A}{L(s)} \right\}, \tag{3.13}$$

where

$$(\pi)_f^*(t,c,s) = \frac{\lambda(t)^2 + L(s)\theta(t,c,s) - \sqrt{(\lambda(t)^2 - L(s)\theta(t,c,s))^2 + 4L(s)^2\lambda(t)^2 h(c,s)}}{2L(s)\lambda(t)^2}.$$

For technical reason of viscosity solution representation proof, we assume that $\pi \leq 1 - \epsilon_A$ in the previous sections. However, to guarantee that the wealth process to be non-negative, we only need the condition that $\pi \leq \frac{1-\epsilon_A}{L(s)}$. Therefore, we release the boundary condition of optimal control $(\pi)^*(t,c,s)$ in the numerical test.

In the case where $\lambda(t)^2 \geq L(s)\theta(t,c,s)$, as $h \to 0$, $(\pi)^*(t,c,s) \to \min\left\{\frac{\theta(t,c,s)}{\lambda(t)^2}, \frac{1-\epsilon_A}{L(s)}\right\}$ which is the optimal control of a standard Merton's model.

3.9.2 Sensitivity of optimal strategies to parameters

Given the closed-form optimal strategies (3.13), we can analyse the sensitivity of optimal strategies to parameters. We test 4 parameters in this section – L, h, θ and λ. Among them, L and h contain the information of conversion risk, while θ and λ represent the information of market risk. Note that θ also depends on the other three parameters:

$$\theta = \frac{\lambda}{\sigma}(\mu - r) + \left(L - \frac{\lambda}{\sigma}L^S\right)(1 + \kappa)h.$$

Despite of the high coupon, the CoCo bond as a fixed income product, is featured to have low volatility compared to the corresponding equity share. Table 3.2 summarizes the volatilities of CoCo bond and equity share of the 5 CoCo contracts in Table 3.1. It is clear that the equity share has the annual volatility range between 20% to 40%, while the CoCo bond volatility range is between 5% to 15%.

Issuing bank	UBS	HSBC	LLOYDS	CreditSuisse	Barclays
CoCo vol	0.05	0.11	0.10	0.08	0.13
Equity vol	0.27	0.21	0.36	0.32	0.30

Table 3.2: Volatility of CoCo bond and equity share

Based on the historical data analysis, we set $\epsilon_A = 0.01, \lambda = 0.1, L = 0.4, r = 0.02, h = 0.1, \mu = 0.1, \sigma = 0.3, L^S = 0.3, \kappa = 2.5$ as benchmark. Under the chosen benchmark parameter set, the value of θ is equal to 0.13. The rationale of benchmark θ value can be seen in Remark 3.5.2. Note that the benchmark parameter set fits a typical EC-CoCo. For a WD-CoCo, L is the written-down fraction $1 - R$ and $L^S = 0$.

We implement the sensitivity test by tuning one of the parameters while keeping the others as the benchmark values.
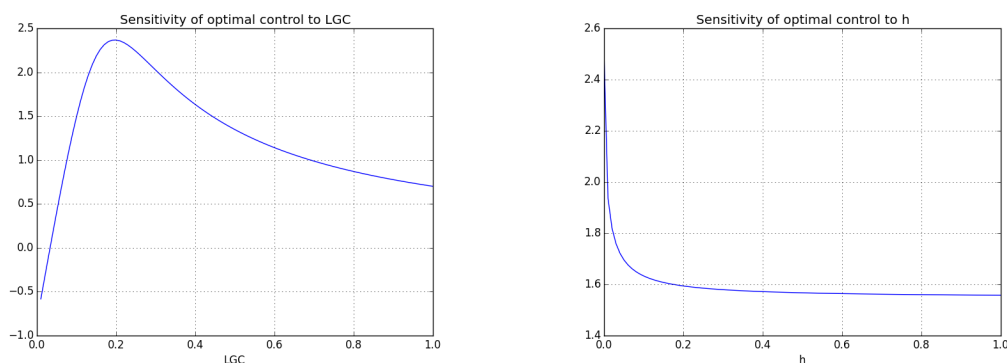
Figure 3.1: Sensitivity of optimal control to LGC and h

Figure 3.1 shows the sensitivity of optimal strategy to parameters L and h. Left panel in Figure 3.1 shows that as L increases, the optimal fraction into CoCo bond increases and then decreases. The humped shape is caused by the fact that the increase of L dominates the drift term of CoCo bond when L is small. However, when L becomes large, it dominates the potential loss at conversion. Right panel in Figure 3.1 shows that as h increases, the optimal strategy is always decreasing. However, the drift of CoCo bond is also increasing with h, thus the decreasing speed of optimal control is very slow when h is large.

Figure 3.2 shows the sensitivity of optimal strategy to parameters θ and λ. θ is the exceed return (including coupon) of CoCo bond to risk-free return, thus the increase of θ results in the increase of optimal control, as shown in the left panel of Figure 3.2. The right panel shows the sensitivity of λ. It is clear that the optimal strategy
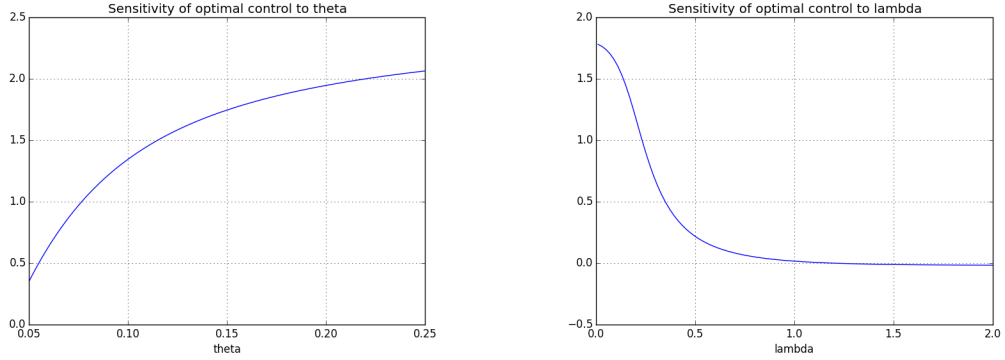
Figure 3.2: Sensitivity of optimal control to θ and λ

decreases with λ, which is the same as the standard Merton's model. However, as λ also occurs in the drift term of CoCo bond, the optimal control goes down and limits to some value as λ increases.

From Figure 3.1 and 3.2, we can find the trade off effect between the conversion risk and yield of CoCo bond. In general, the higher the conversion risk CoCo holder takes, the higher yield they are going to receive. As a result, the optimal strategy of CoCo bond seems always staying positive (no short-selling) with our chosen benchmark parameters. This phenomenon is significantly different from the optimal strategy of equity shares, see Chapter 1.

3.9.3 Performance comparison between investing into CoCo bond and equity

In this section, we compare the performance between investing into CoCo bond and the stock issued by the same bank. We simply assume that the intensity function is only a function of the stock price s and corrected by the coupon rate c. The intensity function h is given by

$$h(c,s) = \min\left\{\max\left\{0, ch_0(s^{-a} - (s^*)^{-a}), 0\right\}, h_M\right\},\tag{3.14}$$

96

where h_0 is a scaling parameter and h_M is the maximum intensity. The coupon c is used as a correction term. The idea is that the larger the coupon c is, the higher the conversion risk is implied by the market. In the meanwhile, the yield $r + \theta(t, c, s)$ is higher as a compensation.

We also assume that the risk premium parameter κ is a function of coupon rate c. Inspired by empirical evidence that the general risk premium ($1 + \kappa$ in our case) lies between 2 and 5, and decreases with conversion risk, we set

$$\kappa(c) = \kappa_0 e^{-c}.$$

where $\kappa_0 = 2.5$ in the benchmark case.

In this comparison test, we choose the following benchmark values:

$$a = 0.8, s^* = 120, h_0 = 300, h_M = 10, c = 0.1.$$

Note that $s^* = 120$ means there is no conversion risk if the stock price is above 120 (we assume stock price starts from 100).

We compare two investors who want to invest into the same company/bank. The company/bank issues common equity (stock) and CoCo bond. One investor invests into CoCo bond and risk-free bank account, and the other invests into stock and bank account. The investor who decides to invest into CoCo bond will follow strategy (3.13). For the equity investor, the optimal strategy can be derived similarly as in Chapter 1:

$$(\pi)^*(c, s) = \min \left\{ (\pi)_f^*(c, s), \frac{1 - \epsilon_A}{L^S(s)} \right\},$$

where

$$(\pi)_f^*(c, s) = \frac{\sigma^2 + L^S(s)(\mu - r) - \sqrt{(\sigma^2 - L^S(s)(\mu - r))^2 + 4L^S(s)^2 \sigma^2 h(c, s)}}{2L^S(s)\sigma^2}.$$

After conversion, $(\pi)^* = (\mu - r)/\sigma^2$ is the solution of the standard Merton's problem.

We will compare the performance between CoCo bond and equity investment based on the above benchmark parameter values and we further assume that:

$$T = 1, S_0 = 100, P_0 = 100, x_0 = 100.$$

We compare EC-CoCo and WD-CoCo respectively.

Equity-converted CoCo bond

Following the general market convention, we assume the EC-CoCo has a fixed conversion price S_F, thus the drop of stock price at conversion is also fixed

$$L_{EC}^S = \frac{\alpha}{\alpha + S_F},$$

and the LGC is given by $LGC = 1 - \frac{S_{\tau-}}{\alpha + S_F}$. By our assumption, LGC is a function of S_t, thus we set

$$L(s) = 1 - \frac{\min\{s, s^*\}}{\alpha + S_F}.$$

The intuition is that with no information about $S_{\tau-}$, we simply estimate $S_{\tau-}|\mathcal{F}_t$ by the current stock price S_t capped by s^*. We further assume that:

$$S_F = 120, \alpha = 50.$$

Under the above setting, $L(s)$ is always greater than $L^S(s)$. It makes sense as the loss of CoCo holder should be more than that of equity holder when conversion occurs.

Figure 3.3 shows the benchmark intensity function with respect to stock prices and the optimal controls of CoCo bond and equity with respect to stock prices. From Figure 3.3, it is clear that both the optimal strategies of CoCo bond and equity increase with the stock price. However, the overall optimal control of CoCo bond is
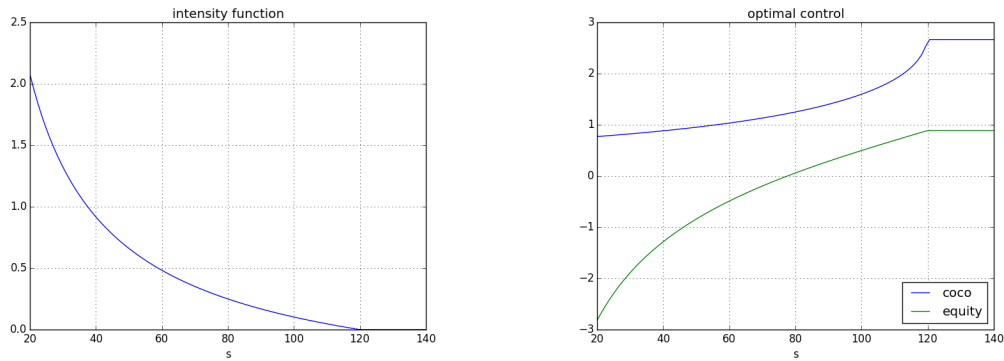
Figure 3.3: Benchmark intensity function and optimal controls (Fully converted case)

much larger than that of equity. This is due to the CoCo bond feature of high yield and low volatility, although CoCo bond is exposed to larger loss when conversion occurs. It is also interesting to notice that when stock price is very low (conversion intensity is very high), the optimal strategy of CoCo bond is still longing. The difference of such behavior between CoCo bond and equity is in that the CoCo bond yield also increases with conversion intensity. As shown in sensitivity analysis, there is a trade off between the high yield and high conversion risk. Note that under the benchmark parameters choice, the initial conversion intensity is 0.1 and the initial CoCo yield $r + \theta$ is 0.15.

We simulate 10000 paths of stock prices and corresponding CoCo bond prices, among which there are 1026 paths containing conversion. Figure 3.4 shows one of the sample paths of stock prices, optimal strategies, and optimal wealth with two different trading assets. The left panel shows price paths of stock S and CoCo bond P. Since both assets are driven by the same Brownian motion, their paths have similar trend but different volatility. At conversion time, CoCo bond price P drops to zero and stock price S jumps down then continues. The middle panel shows the optimal control processes. The overall optimal holding of CoCo bond is much larger than that of equity. The optimal strategy of equity becomes short selling when stock price is low. The right panel shows the sample wealth paths with different trading assets. As the holding of equity is much less than that of CoCo bond, the loss of

equity holder is limited when conversion occurs, while the loss of CoCo bond holders is huge.
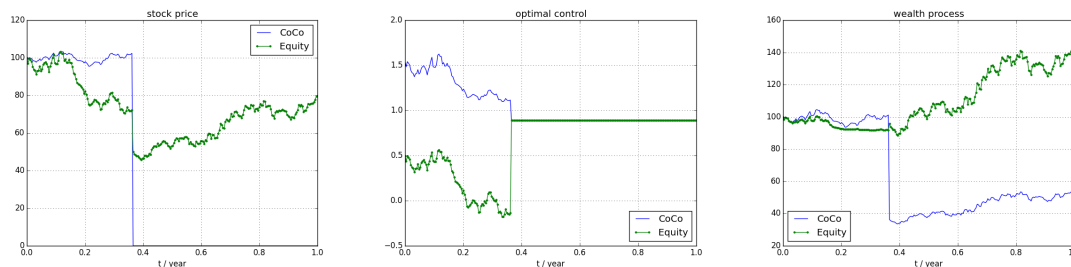
Figure 3.4: Sample paths of stock price, optimal control, and wealth (Fully converted CoCo)

Figure 3.5 shows the statistical distributions of the wealth at time T. The terminal wealth is generated by two assets separately: one is optimal strategy with CoCo bond, the other is optimal strategy with equity. The mid and right panels show the histograms of terminal wealth of these two strategies when conversion indeed occur or does not occur, respectively. It is clear that the overall distributions have marked differences at tail parts, that is, probabilities of great under-performance and out-performance are both higher when investing into CoCo bond. This can be seen in the middle and right panels. When conversion indeed occurs, the performance of investing into CoCo bond is much worse than investing into equity. This is because the optimal strategy of CoCo bond is longing due to the high yield and low volatility, while that of equity becomes short selling when stock price is sufficiently low. However, the right panel shows that when conversion does not occur, the loss probability of investing into CoCo are less than equity (left tail), while the gain probability of investing into CoCo is larger than equity (right tail). This means investing into CoCo is a much better choice if conversion does not occur, due to the benefit of high yield and low volatility.

The terminal distribution statistics are summarized in Table 3.3. From Table 3.3, when conversion occurs, the mean of final wealth with CoCo investment (41.92) is much smaller than that with equity investment (92.68). Meanwhile, both the 2.3% quantile and 97.7% quantile are smaller than investing into equity. However, when

100
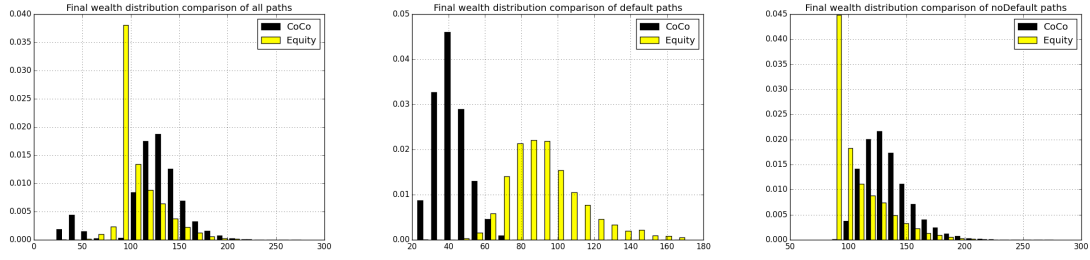
Figure 3.5: Terminal wealth distribution of EC-CoCo

conversion does not occur, the mean of CoCo investment (133.02) is significantly higher than investing into equity (107.48), and the standard deviation (20.47) is smaller than investing into equity (23.02), which means investing into CoCo has more profit and less market risk when conversion does not occur. Therefore, we might conclude that investing into CoCo bond is better choice if the investor has the anticipation that the bank has sufficient capital and the conversion probability is very small.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + CoCo	123.68	33.88	34.99	181.52
All samples + Equity	105.96	23.18	77.29	168.54
Conversion + CoCo	41.92	8.80	26.47	61.93
Conversion + Equity	92.68	20.15	62.20	143.60
No-Conversion + CoCo	133.02	20.47	102.75	183.50
No-Conversion + Equity	107.48	23.02	87.30	170.60

Table 3.3: Sample means, standard deviations, and quantile values (EC-CoCo)

The histograms and tables seem to indicate, for CoCo bond holders, the risk is dominated by conversion. When conversion does not occur, investing into CoCo bond is better than investing into equity in terms of mean and standard deviation of terminal wealth distribution. However, the CoCo bond holders would bear huge loss if the conversion occurred. On the contrary, investing into equity is more sensitive to the market risk, instead of conversion risk.

Written-down CoCo bond

For the WD-CoCo, there is no dilution effect thus $L^S = 0$ and $L(s) = 1 - R$. We assume $R = 0.5$ as the benchmark case. The yield of WD-CoCo is generally larger than that of EC-CoCo, as the loss of WD-CoCo is very likely to be larger. In fact, under the benchmark parameters, the yield of WD-CoCo is equal to 0.2 which is even 5% beyond the yield of EC-CoCo.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + CoCo	125.26	35.26	29.06	177.85
All samples + Equity	109.23	29.71	62.68	174.56
Conversion + CoCo	34.79	7.46	21.92	51.22
Conversion + Equity	90.86	21.83	55.65	145.09
No-Conversion + CoCo	135.61	18.30	107.44	179.16
No-Conversion + Equity	111.33	29.77	64.67	175.48

Table 3.4: Sample means, standard deviations, and quantile values (WD-CoCo)

Table 3.4 shows the statistics of terminal wealth distribution when the CoCo bond is written down at conversion. From Table 3.4, although the loss feature is different, the general conclusion drawn is similar to the case of EC-CoCo.

3.9.4 SENSITIVITY ANALYSIS OF TERMINAL WEALTH DISTRIBUTION

The conclusion drawn in the comparison test is particularly on the benchmark parameter set. We investigate the sensitivity of terminal wealth distribution to the market parameters in this section. Throughout this section, we assume that the CoCo bond is equity-converted type (EC-CoCo).

Sensitivity to Coupon Rate

The coupon rate c plays an important role in our model setting. On one hand, the yield of CoCo bond $r + \theta$ is impacted by coupon rate. Intuitively, the larger the coupon rate is, the higher the CoCo yield is. On the other hand, a large coupon reflects a high conversion probability. Therefore it is a trade off between chasing the

high yield and avoiding the high conversion risk. We investigate the sensitivity of terminal wealth distribution to the coupon rate c. We consider two scenarios in this section: 1) high coupon rate $c = 0.2$ and 2) low coupon rate $c = 0.05$. The other parameters are assumed to be the same as benchmark case.

(1) High coupon rate.

We simulate 10000 paths of stock prices and corresponding CoCo bond prices, among which there are 1822 paths containing conversion. Compared with the benchmark case, the number of converted paths is significantly increased due to the high conversion intensity. In fact, under the scenario $c = 0.2$, the initial conversion intensity is 0.2 and the initial yield of CoCo bond is 0.24.

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	138.04 (11.61%)	54.69 (61.42%)	36.79 (5.14%)	256.52 (41.32%)
All (Equity)	106.04 (0.08%)	25.04 (8.02%)	79.93 (3.42%)	172.26 (2.21%)
C (CoCo)	52.88 (26.15%)	15.24 (73.18%)	29.47 (3.42%)	92.46 (2.21%)
C (Equity)	112.36 (21.23%)	35.86 (77.93%)	65.70 (5.63%)	201.06 (40.01%)
NC (CoCo)	157.01 (18.03%)	40.37 (97.22%)	108.09 (5.20%)	269.87 (47.07%)
NC (Equity)	104.64 (-2.64%)	21.66 (-5.91%)	85.84 (-1.67%)	164.36 (-3.66%)

Table 3.5: Sample means, standard deviations, and quantile values (High coupon rate).

Table 3.5 shows the statistics of terminal wealth distribution with $c = 0.2$. Note that we use C to denote conversion and NC to denote no conversion. The percentage in the bracket shows the relative change compared with the benchmark case. Due to the high conversion risk, the trading strategy of equity becomes more conservative, thus the mean of equity terminal wealth when conversion occurs increases, while that when conversion does not occur decreases. However, the overall performance of CoCo investment (in any scenario) shows significant improvement in mean and percentiles (thus the whole distribution) compared to the benchmark case. Although the high coupon implies the high conversion risk, the high yield compensates for it. As a result, the performance of CoCo investment is much better than equity when conversion does not occur, while the opposite phenomenon happens when conversion

occurs.

<u>(2) Low coupon rate.</u>

In this scenario, we assume $c = 0.05$, thus the initial conversion intensity is 0.05 and initial yield of CoCo bond is slightly higher than 0.1. We simulate 10000 paths of stock prices and corresponding CoCo bond prices, among which there are 526 paths containing conversion.

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	116.01 (-6.20%)	26.72 (-21.13%)	33.73 (-3.60%)	172.80 (-4.80%)
All (Equity)	106.95 (0.93%)	25.67 (10.74%)	75.85 (-1.86%)	171.05 (1.49%)
C (CoCo)	35.13 (-16.20%)	6.96 (-20.91%)	51.74 (95.47%)	22.29 (-64.01%)
C (Equity)	80.47 (-13.17%)	16.31 (-19.06%)	53.44 (-14.08%)	122.77 (-14.51%)
NC (CoCo)	120.50 (-9.41%)	19.17 (-6.35%)	97.66 (-4.95%)	173.88 (-5.24%)
NC (Equity)	108.42 (0.87%)	25.29 (9.86%)	79.72 (-8.68%)	172.10 (0.88%)

Table 3.6: Sample means, standard deviations, and quantile values (Low coupon rate).

Table 3.6 shows the statistics of terminal wealth distribution with $c = 0.05$. From Table 3.6, although the mean of CoCo terminal wealth is still larger than equity when conversion does not occur, the difference becomes much smaller.

Table 3.5 and Table 3.6 indicate that, the larger the coupon rate c is, the better performance the CoCo investment can achieve, especially when no conversion occurs.

Sensitivity to Risk Premium

The coupon rate determines both the CoCo yield and conversion risk, while the risk premium parameter κ_0 only contributes to the yield of CoCo bond. We investigate the sensitivity of terminal wealth distribution to the κ_0 in this section, thus decouple the impact of yield to the terminal wealth distribution. We also consider two scenarios: 1) high risk premium with $\kappa_0 = 5$ and 2) low risk premium with $\kappa_0 = 1$.

<u>(1) High risk premium.</u>

Under the scenario $\kappa_0 = 5$, the initial yield of CoCo bond is 0.22. Table 3.7 shows the

statistics of terminal wealth distribution. It is clear that the equity terminal wealth is not sensitive to the risk premium and the slight difference from benchmark case comes from simulation error. For the CoCo investment, the mean of no-conversion CoCo terminal wealth distribution is much increased, thus a higher risk premium increases the profit of CoCo holder significantly when no conversion occurs. Of course, since high risk premium does not increase the conversion risk, the trading strategy is relatively more aggressive, thus the loss is even larger when conversion does occur.

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	157.51 (27.35%)	66.44 (96.10%)	23.98 (-31.47%)	307.53 (69.42%)
All (Equity)	105.89 (-0.07%)	23.12 (-0.26%)	78.12 (1.07%)	167.07 (-0.87%)
C (CoCo)	30.62 (-26.96%)	9.77 (11.02%)	15.97 (-39.67%)	55.32 (-10.67%)
C (Equity)	92.02 (-0.71%)	20.37 (1.09%)	57.86 (-6.98%)	140.45 (-2.19%)
NC (CoCo)	171.47 (28.91%)	54.06 (164.09%)	112.90 (9.88%)	312.84 (70.49%)
NC (Equity)	107.41 (-0.07%)	22.90 (-0.52%)	87.33 (0.03%)	168.74 (-1.09%)

Table 3.7: Sample means, standard deviations, and quantile values (High risk premium).

(2) Low risk premium.

Under the scenario $\kappa_0 = 1$, the initial yield of CoCo bond is 0.11. The corresponding terminal wealth distribution statistics are shown in Table 3.8.

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	110.65 (-10.54%)	24.28 (-28.34%)	52.12 (48.96%)	165.84 (-8.64%)
All (Equity)	106.27 (0.29%)	23.35 (0.73%)	79.31 (2.61%)	168.39 (-0.09%)
C (CoCo)	60.91 (45.30%)	12.48 (41.82%)	37.81 (42.84%)	90.09 (45.47%)
C (Equity)	93.72 (1.12%)	20.92 (3.82%)	60.01 (-3.52%)	146.09 (1.73%)
NC (CoCo)	116.28 (-12.58%)	18.08 (-11.68%)	95.93 (-6.64%)	167.67 (-8.63%)
NC (Equity)	107.69 (0.20%)	23.19 (0.74%)	87.26 (-0.05%)	169.72 (-0.52%)

Table 3.8: Sample means, standard deviations, and quantile values (Low risk premium).

Table 3.8 confirms that the CoCo bond is not as attractive as that with high risk premium. Although it is still true that the performance of CoCo investment is better

than equity in terms of mean and standard deviation, the difference between CoCo and equity is not as significant as before. However, when conversion does occur, the loss of CoCo investment is much more than that of equity investment.

Sensitivity to Initial Stock Price

Under the benchmark case, we assume that $S_0 = 100$ such that the initial conversion intensity is around 0.1. However, under the financial distress period (e.g. 2008-2009 financial crisis period), the conversion intensity could be much higher than the benchmark value. We test in this section the case where the initial stock price is low, which equivalently means the initial conversion probability is large. We set $S_0 = 80$. The corresponding initial conversion intensity becomes 0.25, and thus the yield of CoCo bond is 0.4. We simulate 10000 paths among which there are 2584 containing conversion. The conversion probability is close to one quarter.

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	148.30 (19.91%)	57.64 (70.13%)	34.16 (-2.37%)	244.76 (34.84%)
All (Equity)	104.07 (-1.78%)	17.91 (-22.74%)	79.79 (3.23%)	156.21 (-7.32%)
C (CoCo)	49.85 (18.92%)	13.04 (48.18%)	27.24 (2.91%)	76.48 (23.49%)
C (Equity)	107.18 (15.65%)	26.04 (29.23%)	67.19 (8.02%)	168.82 (17.56%)
NC (CoCo)	174.22 (30.97%)	30.48 (48.90%)	129.19 (25.73%)	249.75 (36.10%)
NC (Equity)	103.25 (-3.94%)	14.95 (-35.06%)	92.48 (5.93%)	147.73 (-13.41%)

Table 3.9: Sample means, standard deviations, and quantile values (Low initial stock price).

Table 3.9 shows the statistics of terminal wealth distribution when initial stock price $S_0 = 80$. Due to the feature of high yield, the mean of CoCo terminal wealth when conversion does not occur is much higher than that of equity terminal wealth. Compared with Table 3.7, it is interesting to find that the means of CoCo terminal wealth under no-conversion case are similar, however, the standard deviation is much smaller. As discussed, increasing the risk premium has no impact to the conversion risk, thus the trading strategy is much more aggressive. As a consequence, the mean of CoCo terminal wealth under conversion case in Table 3.7 is significantly lower than that in Table 3.9.

Sensitivity to Written-Down Fraction

For a WD-CoCo, the written-down fraction $1 - R$ is an important feature which determines the loss of CoCo holders when conversion occurs. We analyze the terminal wealth distribution sensitivity to R in this section. We also consider two scenarios in this section: 1) high written-down fraction $R = 0.2$ and 2) low written-down fraction $R = 0.8$.

(1) High written-down fraction.

In this scenario, we assume $R = 0.2$. Table 3.10 shows the statistics of terminal wealth distribution. Compared with Table 3.4, it is clear that the performance of equity investment is similar. It is interesting to note that the performance of CoCo investment is also similar to the benchmark case. The high written-down fraction has multiple impacts. The yield increases with the written-down fraction and the optimal control becomes more conservative. The combined effects seem to compensate for each other, and result in the robust performance under different written-down fractions.

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	123.92 (-1.07%)	35.11 (-0.43%)	30.00 (3.23%)	183.61 (3.24%)
All (Equity)	109.05 (-0.16%)	29.24 (-1.58%)	62.34 (-0.54%)	174.25 (-0.18%)
C (CoCo)	36.59 (5.17%)	8.54 (14.48%)	21.50 (-1.92%)	55.70 (8.75%)
C (Equity)	90.18 (-0.75%)	22.75 (4.21%)	55.46 (-0.34%)	144.04 (-0.72%)
NC (CoCo)	134.04 (-1.16%)	19.50 (6.56%)	108.75 (1.22%)	185.80 (3.71%)
NC (Equity)	111.00 (-0.30%)	29.28 (-1.65%)	64.05 (-0.96%)	174.85 (-0.36%)

Table 3.10: Sample means, standard deviations, and quantile values (High written-down fraction).

(2) Low written-down fraction.

In this scenario, we assume $R = 0.8$. Table 3.11 confirms that the performance of CoCo bond investment is robust under different written-down fraction levels.

Table 3.10 and Table 3.11 indicate that the performance of WD-CoCo investment is generally robust with different written-down fraction levels, with slight difference

	mean	std dev	2.3% quantile	97.7% quantile
All (CoCo)	127.36 (1.68%)	39.52 (12.08%)	29.19 (0.45%)	196.15 (10.29%)
All (Equity)	109.65 (0.38%)	29.56 (-0.50%)	62.83 (0.24%)	175.66 (0.63%)
C (CoCo)	36.29 (4.31%)	9.18 (23.06%)	22.64 (3.28%)	56.70 (10.70%)
C (Equity)	90.87 (0.01%)	22.72 (4.08%)	54.35 (-2.34%)	142.73 (-1.63%)
NC (CoCo)	137.72 (1.56%)	26.06 (42.40%)	93.07 (-13.37%)	197.94 (10.48%)
NC (Equity)	111.61 (0.25%)	29.57 (-0.67%)	64.81 (0.22%)	179.83 (2.48%)

Table 3.11: Sample means, standard deviations, and quantile values (Low written-down fraction).

in tail distribution of terminal wealth.

3.9.5 OPTIMAL STRATEGIES FOR POWER UTILITY

For power utility $U(x) = (1/\gamma)x^\gamma$, $0 < \gamma < 1$, the post-conversion case is well known with the optimal control $\pi^S = (\mu^S - r)/(\sigma^2(1 - \gamma))$ and the post-conversion value function $v_1(t, x) = (1/\gamma)x^\gamma g_1(t)$, where

$$g_1(t) = \exp\left(\left(r\gamma + \frac{\gamma}{2(1 - \gamma)}\left(\frac{\mu - r}{\sigma}\right)^2\right)(T - t)\right).$$

We conjecture that the pre-conversion value function takes the form

$$v_0(t, x, s) = \frac{x^\gamma}{\gamma}f(t, s). \tag{3.15}$$

Substituting (3.15) into (3.8), we get a linear PDE for f:

$$\frac{\partial f}{\partial t} + \sup_{\pi \in A}\left\{b(t, s, \pi)f_s + \frac{1}{2}\sigma^2 s^2 f_{ss} - \beta(t, c, s, \pi)f + g(t, c, s, \pi)\right\} = 0 \tag{3.16}$$

with the terminal condition $f(T, s) = 1$, where

$$b(t, s, \pi) := (\mu + \gamma\sigma\lambda(t)\pi)s,$$

$$\beta(t,c,s,\pi) := -r\gamma + h(c,s) - \gamma\theta(t,c,s)\pi - \frac{1}{2}\gamma(\gamma-1)\lambda(t)^2\pi^2,$$

$$g(t,c,s,\pi) := h(c,s)g_1(t)(1-L(s)\pi)^\gamma.$$

By Pham (2009) (Remark 3.4.2), the HJB equation (3.16) can be mapped to the following optimal control problem:

$$dY_t = b(t,Y_t,\pi_t)dt + \sigma Y_t dZ_t, \quad Y_0 = s, \tag{3.17}$$

$$J(t,y,\pi) = \mathbb{E}\left[\int_t^T \Gamma(t,c,u)g(u,c,Y_u,\pi_u)du + \Gamma(t,c,T)\right],$$

$$v(t,s) = \sup_{\pi\in\mathcal{A}} J(t,s,\pi), \tag{3.18}$$

where Z_t is a standard Brownian motion and

$$\Gamma(t,c,u) := \exp\left\{-\int_t^u \beta(l,c,Y_l,\pi_l)dl\right\}.$$

By our theoretical result, we claim that the value function $v(t,s)$ is the unique viscosity solution of the HJB equation (3.16).

We adopt the numerical method in Kushner and Dupuis (2001) to solve the above optimization problem. Kushner and Dupuis approximate the process Y_t by a Markov chain, which transits from Y_t to $Y_{t+\Delta t}$ with probability $Q^{\delta,\Delta t}(Y_t,Y_{t+\Delta t} \mid \pi)$. Note that δ is the step size of space and Δt is the step size of time. Fix the time horizon $[t,T]$ and a time discretization

$$\Delta t := \frac{T-t}{N},$$

where $N \geq 1$ is an integer. For simplicity, we assume $t = 0$. According to Kushner and Dupuis (2001), we approximate the process (3.17) by a Markov chain with

transition probabilities:

$$Q^{\delta,\Delta t}(s, s \mid \pi) := 1 - \frac{\Delta t}{\delta}|b(t, s, \pi)| - \frac{\Delta t}{\delta^2}\sigma^2 s^2$$

$$Q^{\delta,\Delta t}(s, s \pm \delta \mid \pi) := \frac{\Delta t}{\delta}b(t, s, \pi)^{\pm} + \frac{\Delta t}{2\delta^2}\sigma^2 s^2$$

where $x^+ := \max\{x, 0\}$, $x^- := \max\{-x, 0\}$.

The numerical scheme is based on the following backward discretized dynamic programming principle (discretized DPP):

$$v(k\Delta t, Y_{k\Delta t}) \approx \sup_{\pi_k \in A} \mathbb{E}\Bigg[g(k\Delta t, c, Y_{k\Delta t}, \pi_k)\Delta t$$

$$+ \Gamma(k\Delta t, c, (k+1)\Delta t)v\left((k+1)\Delta t, Y_{(k+1)\Delta t}\right)\Bigg],$$

where π_k is the piece-wise constant control and we approximate $\Gamma(k\Delta t, c, (k+1)\Delta t) \approx \exp\{-\beta(k\Delta t, c, Y_{k\Delta t}, \pi_k)\Delta t\}$.

Note that we can apply the similar algorithm to the equity investment case. We apply above algorithm to the benchmark case with $\gamma = 0.5$, $\delta = 5$ and step size of time $\Delta t = 0.1$. The CoCo bond is assumed to be EC-CoCo.

	mean	std dev	2.3% quantile	97.7% quantile
All samples + CoCo	141.75	75.47	11.09	301.97
All samples + Equity	111.66	65.23	28.52	284.05
Conversion + CoCo	21.94	7.86	4.55	35.63
Conversion + Equity	99.42	34.49	67.76	180.24
No-Conversion + CoCo	157.22	60.99	115.75	316.29
No-Conversion + Equity	119.05	63.14	38.50	290.45

Table 3.12: Sample means, standard deviations, and quantile values of power utility (EC-CoCo)

The result is similar to the log utility case such that the main risk of investing into CoCo bond is contained in the conversion. As long as no conversion occurs, investing into CoCo bond has better performance than equity in terms of both mean and

standard deviation of terminal wealth distribution. However, if conversion occurs, the loss of CoCo holders is significantly larger than the equity holders.

3.10 Conclusions

In this chapter we consider a utility maximization problem of CoCo bond. We assume that the conversion intensity is a function of the coupon rate and the stock price of the issuing bank, and the intensity value is zero when stock price is above the implied barrier level. The conversion of CoCo bond induces an immediate drop in the stock price of the issuing bank. We analyse the loss given conversion structure of different CoCo bond contracts and we prove the pre-conversion value function is the unique viscosity solution of the HJB equation. We also compare and analyse the statistical distributions of terminal wealth of log utility and power utility based on two investment assets, one with CoCo bond and the other with stock of the issuing bank. Our simulation results show that, the CoCo bond holders bear much more loss than equity holders when conversion occurs. However, investing into CoCo bond gets more profit (mean) while bearing less market risk (volatility) as long as conversion does not occur. There remain many open questions in utility maximization of CoCo bond, for example, what is the proper CoCo bond price dynamic if the coupon is paid continuously at rate c to the face value instead of the market value. We leave these and other questions to future research.

4

CONCLUSIONS

This thesis is concentrated on the dynamic portfolio optimization with credit risk.

In Chapter 3, we consider a utility maximization problem with defaultable stocks and looping contagion risk. We assume that the default intensity of one company depends on the stock prices of itself and other companies, and the default of the company induces immediate drops in the stock prices of the surviving companies. We prove that the value function is the unique viscosity solution of the HJB equation. By numerical tests, we compare the statistical distributions of the terminal wealth based on two strategies, one using the full information of intensity process and the other a proxy constant intensity process. Our numerical result shows that, statistically, the looping contagion risk model and exogenous factor model have similar performance in general market situations. However, if there are big drops of stock prices at the start of investment, one may greatly improve the performance of investment if one uses the looping contagion risk model.

In Chapter 4, we consider a dynamic portfolio optimization with contingent convertible (CoCo) bond. We model the conversion of CoCo bond by reduced-form approach and assume that the conversion intensity is a deterministic function of the coupon rate and the issuing bank's stock price. Theoretically, we prove that the value function is the unique viscosity solution of the corresponding HJB equation. Practically, we compare the performance between investing into CoCo bond and the issuing bank's stock. We analyse the statistical distributions of terminal wealth of log utility and power utility based on these two investment choices. Our simulation results show that, the CoCo bond holders bear more loss than equity holders if conversion occurs. However, investing into CoCo bond gets more profit (mean) while bearing less market risk (volatility) as long as conversion does not occur.

We develop a numerical method to solve the dynamic portfolio optimization problem with Heston model framework in Appendix C. The numerical method is based on approximating the controlled process by a tree model and applying dynamic programming principle to the approximated tree model. We use numerical test to show the convergence of the numerical method.

There remain many open questions in this thesis. For example, the BSDE simulation method for power utility, the proper CoCo bond price dynamic if the coupon is paid continuously at rate c to the face value instead of the market value, the theoretical proof of the convergence of numerical method. We leave these and other questions to future research.

REFERENCES

[1] Akyildirim, E., Y. Dolinsky and H. Mete Soner (2014), *Approximating Stochastic Volatility by Recombinant Trees*, Ann. Appl. Probab., 24(5), 2176-2205.

[2] Bielecki, T. R. and I. Jang (2006), *Portfolio Optimization with a Defautable Security*, Asia-Pacific Financial Markets, 13, 113-127.

[3] Bielecki, T. R. and M. Rutkowski (2003), *Credit Risk: Modeling, Valuation and Hedging*, Springer.

[4] Bo, L. and A. Capponi (2016), *Optimal Investment in Credit Derivatives Portfolio under Contagion Risk*, Mathematical Finance, 26(4), 785-834.

[5] Bouchard, B. and N. Touzi (2011), *Weak Dynamic Programming Principle for Viscosity Solutions*, SIAM Journal on Control and Optimization, 49 (3), 948-962.

[6] Bo, L., Y. Wang and X. Yang (2010), *An Optimal Portfolio Problem in a Defaultable Market*, Advances in Applied Probability, 42, 689-705.

[7] Brigo, D., J. Garcia and N. Pede (2015), *CoCo Bonds Pricing with Equity- and Credit-Calibrated First Passage Firm Value Models*, Int. J. Theor. Appl. Finance, 18(3).

[8] Brigo, D. and M. Morini (2013), *Counterparty Credit Risk, Collateral and Funding*, Wiley.

[9] Callegaro, G., Jeanblanc, M., and Runggaldier, W. (2012), *Portfolio Optimiza-*

tion in a Defaultable Market under Incomplete Information, Decisions in Economics and Finance, 35(2), 91-111.

[10] Capponi, A. and J. E. Figueroa-Lopez (2011), *Dynamic Portfolio Optimization with a Defaultable Security and Regime Switching*, Mathematical Finance, 24, 207-249.

[11] Capponi, A. and C. Frei (2017), *Systemic Influences on Optimal Equity-Credit Investment*, Management Science, 63, 2756-2771.

[12] Chung, T.K. and Y.K. Kwok (2015), *Enhanced Equity-Credit Modeling for Contingent Convertibles*, Quantitative Finance, 16(10), 15111527.

[13] Cheridito, P. , H. M. Soner , N. Touzi and N. Victoir (2007), *Second Order Backward Stochastic Differential Equations and Fully non-linear Parabolic PDEs*, Communications on Pure and Applied Mathematics, 60, 1081-1110.

[14] Crandall, M., H. Ishii and P.L. Lions (1992), *User's Guide to Viscosity Solutions of Second Order Partial Differential Equations*, Bulletin of the American Mathematical Society, 27, 1-67.

[15] Corcuera, J.M., J.D. Spiegeleer, A. Ferreiro-Castilla, A.E. Kyprianou, D.B. Madan and W. Schoutens (2013), *Pricing of Contingent Convertibles under Smile Conform Models*, J. Credit Risk, 9(3), 121-140.

[16] Cheridito, P. and Z. Xu (2015), *A Reduced Form CoCo Model with Deterministic Conversion Intensity*, J. Risk, 17(3), 118.

[17] Davis, M. H. and A.R. Norman (1990), *Portfolio Selection with Transaction Costs*, Mathematics of Operations Research, 15(4), 676-713.

[18] Duffie, D. and K. Singleton (1999), *Modeling Term Structures of Defaultable Bonds*, Rev. Financ. Stud., 12, 687720.

[19] Flemming, H. and D. Hernandez-Hernandez (2003), *An Optimal Consumption Model with Stochastic Volatility*, Finance and Stochastics, 7(2), 245-262.

[20] Fouque, J.P., R. Sircar and T. Zariphopoulou (2017), *Portfolio Optimization and Stochastic Volatility Asymptotics*, Mathematical Finance, 27(3), 704-745.

[21] Gu, J.W., W.K. Ching, T.K. Siu, and H. Zheng (2013), *On Pricing Basket Credit Default Swaps*, Quantitative Finance, 13, 1845-1854.

[22] Heynderickx, W., J. Cariboni, W. Schoutens and B. Smits (2016), *The Relationship between Risk-neutral and Actual Default Probabilities: the Credit Risk Premium*, Applied Economics, 48:42, 4066-4081.

[23] Hou, Y. and X. Jin (2002), *Optimal Investment with Default Risk*, FAME Research Paper No. 46, Switzerland.

[24] Jarrow, R. and F. Yu (2001), *Counterparty Risk and the Pricing of Defaultable Securities*, Journal of Finance, 56, 1765-1799.

[25] Jarrow, R. , D. Lando and F. Yu (2005), *Default Risk and Diversification: Theory and Applications*, Mathematical Finance, 15, 1-26.

[26] Jang, H.J., Y.H. Na, H. Zheng (2018), *Contingent Convertible Bonds with the Default Risk Premium*, International Review of Financial Analysis, 59, 77-93.

[27] Janecek, K. and S.E. Shreve (2004), *Asymptotic Analysis for Optimal Investment and Consumption with Transaction Costs*, Finance and Stochastics, 8(2), 181-206.

[28] Jiao, Y. and H. Pham (2011), *Optimal Investment with Counterparty Risk: a Default-density Modeling Approach*, Finance and Stochastics, 15, 725-753.

[29] Jiao, Y. and H. Pham (2013), *Optimal Investment under Multiple Defaults Risk: a BSDE-decomposition Approach*, Annals of Applied Probability, 23, 455-491.

[30] Karatzas, I. and S.E. Shreve (2000), *Brownian Motion and Stochastic Calculus*, Springer.

[31] Kraft, H. (2005), *Optimal Portfolios and Heston's Stochastic Volatility Model: an Explicit Solution for Power Utility*, Quantitative Finance, 5(3), 303-313.

[32] Kraft, H. (2009), *Optimal Portfolios with Stochastic Short Rate: Pitfalls when the Short Rate is Non-Gaussian or the Market Price of Risk is Unbounded*, International Journal of Theoretical and Applied Finance, 12(6), 767-796.

[33] Korn, R. and H. Kraft (2002), *A Stochastic Control Approach to Portfolio Problems with Stochastic Interest Rates*, SIAM J. Control Optim., 40(4), 1250-1269.

[34] Korn, R. and H. Kraft (2003), *Optimal Portfolios with Defaultable Securities: a Firm Value Approach*, International Journal of Theoretical and Applied Finance, 6, 793-819.

[35] Kushner, H.J. and P. Dupuis (2001), *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer.

[36] Lakner, P. and W. Liang (2008), *Optimal Investment in a Defaultable Bond*, Mathematics and Financial Economics, 1, 283-310.

[37] Leung, C.M. and Y.K. Kwok (2015), *Numerical Pricing of CoCo Bonds with Parisian Trigger Feature Using the Fortet Method*, Working Paper, Hong Kong Baptist University.

[38] Liberman, G. (1996), *Second Order Partial Differential Equations*.

[39] Liu, C. and H. Zheng (2016), *Asymptotic Analysis for Target Asset Portfolio Allocation with Small Transaction Costs*, Insurance: Mathematics and Economics, 66, 59-68.

[40] Markowitz, H. (1952), *Portfolio Selection*, The journal of Finance, 7(1), 77-91.

[41] Ma. J.T., D.Y. Deng and H. Zheng (2016), *Dual Control Monte-Carlo Method for Tight Bounds of Value Function in Regime Switching Utility Maximization*, European J Operational Research, 262, 851-862.

[42] Merton, R.C. (1969), *Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case*, The Review of Economics and Statistics, 51(3), 247-257.

[43] Merton, R.C. (1971), *Optimum Consumption and Portfolio Rules in a Continuous-time Model*, Journal of Economic Theory, 3), 373-413.

[44] Noh, E.J. and J.H. Kim (2011), *An Optimal Portfolio Model with Stochastic Volatility and Stochastic Interest Rate*, Journal of Mathematical Analysis and Applications, 375(2), 510-522.

[45] Øksendal, B. (2003), *Stochastic Differential Equations*, Springer.

[46] Pham, H. (2009), *Continuous-Time Stochastic Control and Optimization with Financial Applications*, Springer.

[47] Pham, H. (2002), *Smooth Solutions to Optimal Investment Models with Stochastic Volatilities and Portfolio Constraints*, Applied Mathematics and Optimization, 46, 55-78.

[48] Protter, P. E. (2005), *Stochastic Integration and Differential Equations*, Springer.

[49] Rogers, L.C.G. (2013), *Optimal Investment*, Springer.

[50] Spiegeleer, J.D. and W. Schoutens (2012), *Pricing Contingent Convertibles: A Derivatives Approach*, J. Derivatives, 20(2), 2736.

A

EMPIRICAL EVIDENCE OF

CONTAGION RISK

We give simple empirical example in this appendix, that shows the evidence such
that there exists looping contagion risk in the financial market.
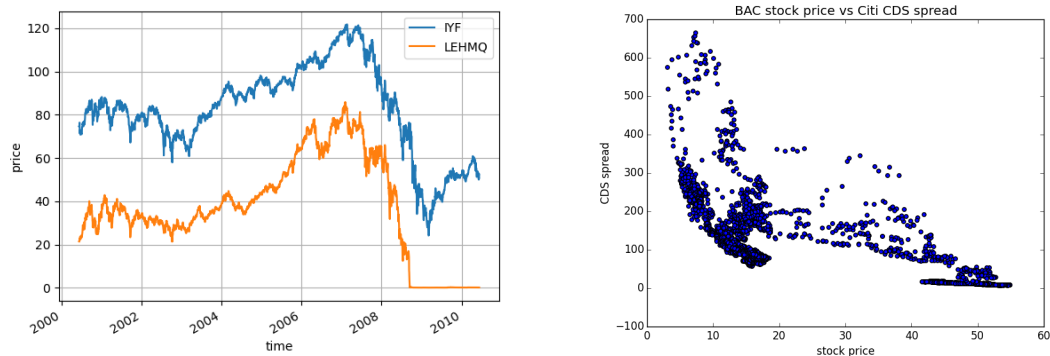
Figure A.1: Looping contagion risk existence evidence (left: comparison of iShares US Financials
ETF (IYF) and Lehman Brothers (LEHMQ); right: comparison of Citibank CDS spread and Bank of
America stock price). Data Source: Bloomberg

The left panel of Figure A.1 shows the comparison of Lehman Brothers stock price and the iShares US Financial ETF (IYF) price. It is clear that there is strong correlation between the stock prices of Lehman Brothers and IYF. The price of IYF can be treated as a signal or factor that reflects the entire situation of financial sector. On one hand, it implies that the default risk of any single firm in financial sector goes up when IYF price drops down. On the other hand, the failure of any financial institutions (e.g. Lehman Brothers) leads to sharp falls in IYF price.

The right panel of Figure A.1 shows the relation between Bank of America stock price and Citi CDS spread. The CDS spread reflects the default probability from market point of view. From Figure A.1, there exists clear pattern such that the CDS spread of Citibank decreases with the stock price of Bank of America. The similar pattern can be seen among the other financial institutions, e.g. Barclays, JP Morgan, Goldman Sachs.

Therefore, Figure A.1 is an empirical evidence which indicates that there exists strong contagion risk in the financial market.

B

ONE-SIDED CONTAGION RISK

We show some extra interesting results of the one-sided contagion risk defined in Example 2.2.1. Due to the simple structure that there is only one default in the model setting, the problem can be naturally split into pre-default case and post-default case. The latter is a standard utility maximization problem as stock P disappears and the post-default value function v_1 is a function of time t and wealth x only, see Pham (2009). We will only discuss the pre-default case in this section. The pre-default value function is defined by

$$v_0(t, x, s, p) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^\pi)|X_t = x, S_t = s, P_t = p, H_t = 0\right]$$

for $(t, x, s, p) \in [0, T] \times (0, \infty)^3$.

As shown in Theorem 2.4.2, the pre-default value function v_0 is continuous in $(t, x, s, p) \in [0, T] \times [0, \infty) \times (0, \infty)^2$. It is clear that the natural boundary condition at $x = 0$ is $v_0(t, 0, s, p) = U(0)$. The natural boundary condition at $s = 0$ is

the optimal value function in a market with two securities, one defaultable stock P and one riskless bank account B. However, due to the lack of continuity at $s = 0$, the natural boundary condition may be irrelevant. Same for the boundary condition for $p = 0$. To compensate for this, we make the following assumption for the pre-default value function.

Assumption B.0.1. Denote the boundary of $[0, \infty)^2$ by $\mathcal{B}$, then $\mathcal{B} := \{\{0\} \times [0, \infty)\} \cup \{[0, \infty) \times \{0\}\}$. Assume $\hat{v}_0(t, x, s, p) = \lim_{(t', x', s', p') \to (t, x, s, p)} v_0(t', x', s', p')$ exists for any $(t, x, s, p) \in [0, T] \times [0, \infty) \times \mathcal{B}$.

By definition, $\hat{v}_0$ is a proper boundary condition for the pre-default value function v_0 at $s = 0$ or $p = 0$, which equals the natural boundary condition if v_0 is continuous at $s = 0$ and $p = 0$. Define

$$
\tilde{v}_0(t, x, s, p) = \begin{cases} v_0(t, x, s, p) & \text{if } (s, p) \notin \mathcal{B} \\ \hat{v}_0(t, x, s, p) & \text{if } (s, p) \in \mathcal{B}. \end{cases}
$$

The modified pre-default value function $\tilde{v}_0$ is continuous on $[0, T] \times [0, \infty)^3$ by Theorem 2.4.2 and Assumption B.0.1 and satisfies the following HJB equation by the DPP

$$
- \sup_{\pi \in A} \mathcal{L}^\pi w(t, x, s, p) = 0 \tag{B.1}
$$

for $(t, x, s, p) \in [0, T) \times (0, \infty)^3$ with terminal condition $w(T, x, s, p) = U(x)$, where $\mathcal{L}^\pi$ is the infinitesimal generator of processes S, P and X with control π, given by

$$
\begin{aligned}
\mathcal{L}^\pi w(t, x, s, p) = {}& \frac{\partial w}{\partial t} + (r + \theta^T \pi) x w_x + \mu^S s w_s + \mu^P p w_p + \frac{1}{2} \pi^T \Sigma \pi x^2 w_{xx} \\
& + \frac{1}{2} (\sigma^S s)^2 w_{ss} + \frac{1}{2} (\sigma^P p)^2 w_{pp} + \sigma^S m^T \pi x s w_{xs} + \sigma^P n^T \pi x p w_{xp} \\
& + \rho \sigma^S \sigma^P s p w_{sp} + h(s, p)(v_1\left(t, x\left(1 - L^T \pi\right)\right) - w),
\end{aligned}
$$

where v_1 is the post-default value function, $m = (\sigma^S, \rho\sigma^P)^T$, $n = (\rho\sigma^S, \sigma^P)^T$.

By Remark 2.6.5, the structure condition holds for the one-sided contagion model setting. Therefore, all the conditions in Theorem 2.6.6 are satisfied. We have the following viscosity solution properties for the modified pre-default value function.

Theorem B.0.2. *The modified pre-default value function $\tilde{v}_0$ is the unique continuous viscosity solution of (B.1) on $[0, T) \times (0, \infty)^3$, satisfying the growth condition $|\tilde{v}_0(t, x, s, p)| \leq K(1 + x^\gamma)$ for some constant $0 < \gamma < 1$, the terminal condition $\tilde{v}_0(T, x, s, p) = U(x)$, and the boundary conditions $\tilde{v}_0(t, 0, s, p) = U(0)$, $\tilde{v}_0(t, x, s, p) = \hat{v}_0(t, x, s, p)$ for $(s, p) \in \mathcal{B}$.*

Remark B.0.3. Note that for the one-sided contagion risk model, the continuity of modified pre-default value function $\tilde{v}_0$ has been proved in prior. Thus the proofs of Theorem 2.8.3 and 2.6.6 which rely on the upper (lower) semi-continuous envelops can be simplified accordingly.

C

Numerical Method Based on Tree approximation

In this appendix, we investigate the numerical method to solve a dynamic portfolio optimization problem under Heston model framework. We consider an investor who dynamically puts her money into a stock and a risk-free bank account. The stock price is driven by the Heston model. Kraft (2005) shows that the optimal strategy and value function can be solved in closed form, under following conditions:

- The market price of risk of Heston model is equal to a linear function of the volatility process;

- The investor's utility function is power utility.

To the best of our knowledge, there has been no successful attempt to solve the portfolio optimization under Heston model explicitly without assuming above con-

ditions. To solve this problem numerically, one can try the numerical method to solving the HJB equation directly. However, it is known that numerically solving the HJB equation under Heston model setting is difficult due to lack of boundary conditions. Ma et al. (2017) introduces the dual control Monte Carlo method to get tight bounds of value function under this framework. Alternatively, one may apply the numerical method introduced by Kushner and Dupuis (2001). In Chapter 3 and 4, we apply the numerical methods by Kushner and Dupuis (2001) to power utility numerical tests, respectively. By guessing that the value function takes the form (2.15), we find that the resulting controlled process (2.18) of the new control problem (2.17) has the volatility term $\sigma(Y_u)$ independent of control π. In general, Kushner and Dupuis' numerical method is designed for the controlled process in the following form

$$dY_t = b(Y_t, \pi_t) + \sigma(Y_t)dW_t,$$

where the volatility term $\sigma(Y_t)$ of the controlled process is assumed to be independent of control π_t. The idea is to approximate the original controlled process Y_t by an appropriate finite state Markov chain.

For the portfolio optimization problem under Heston model framework, we will show later that the controlled process Y_t has the following form $Y_t := (X_t, V_t)$, where X_t is the wealth process and V_t is the stochastic variance following CIR process. It can be checked that the wealth process X_t contains control in the volatility term, see (C.2). Therefore, the theoretical convergence result proved in Kushner and Dupuis (2001) can not be applied. Moreover, the approximated Markov chain starting from $y := (x, v)$ can only take 9 combinations in the next step – $(x, v), (x \pm \delta, v), (x, v \pm \delta), (x \pm \delta, v \pm \delta), (x \pm \delta, v \mp \delta)$. As the wealth process X_t and volatility process V_t normally have different magnitudes, there is strong stability issue of the numerical method if the same δ is applied as the step size for both x and v.

We extend the approach in Kushner and Dupuis (2001) to develop a numerical method to solve the dynamic portfolio optimization problem under Heston model

framework. Instead of approximating the controlled process by a Markov chain with identical constant step size δ, we approximate the continuous dynamic of the controlled process by a discrete non-combining tree, whose step size depends on the volatility value at each step. We show the convergence of our numerical method, by comparing the numerical result (from our approach) with the theoretical result from Kraft (2005) (under power utility) and the tight bounds estimated from Ma et al. (2017) (under non-HARA utility). The numerical comparison shows that the approximated controls and value function from our numerical method converges to the real solution.

Problem Description

Assume that $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a given probability space with filtration $\mathcal{F}_t$ generated by standard Brownian motions W and Z with correlation coefficient $\rho \in (-1, 1)$ and completed with all P–null sets. There exists one stock in the market whose price is denoted by $\{S_t\}_{t \geq 0}$. The dynamic of S_t follows the Heston model

$$
\begin{aligned}
\frac{dS_t}{S_t} &= \mu dt + \sqrt{V_t} dW_t, \quad S_0 = s, \\
dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dZ_t, \quad V_0 = v,
\end{aligned}
\tag{C.1}
$$

with initial conditions $s, v > 0$, given positive parameters $\mu, \kappa, \theta, \sigma$. By standard argument, we assume that the Feller condition is satisfied, which means $2\kappa\theta > \sigma^2$, thus the Heston model has a unique positive solution.

An investor dynamically allocates proportions $(\pi_t, 1 - \pi_t)$ of the total wealth into the stock and the bank account. The admissible control set $\mathcal{A}$ is the set of control processes π that are progressively measurable with respect to the filtration $(\mathcal{F}_t)$ and $\pi_t \in A$ for $\forall t \in [0, T]$. The set A is defined such that the wealth process is non-negative for $\forall t \in [0, T]$. The bank account $\{B_t\}_{t \geq 0}$ follows

$$
\frac{dB_t}{B_t} = rdt,
$$

where $r > 0$ is the constant risk-free rate. The dynamic of the wealth process $(X_t)_{t \geq 0}$ is given by

$$\frac{dX_t}{X_t} = (r + (\mu - r)\pi_t)\, dt + \pi_t \sqrt{V_t} dW_t. \tag{C.2}$$

Note that the volatility term of the wealth process is stochastic and contains control π in the place.

We consider the utility maximization problem defined by

$$\sup_{\pi \in \mathcal{A}} \mathbb{E}[U(X_T^\pi)],$$

where U is a utility function that is continuous, non-decreasing and concave on $(0, \infty)$. Then the value function is defined by

$$G(t, x, v) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_T^\pi) | X_t = x, V_t = v\right], \tag{C.3}$$

Note that throughout this appendix, we denote the value function by G, as the notation v denotes the value of the initial variance process in this appendix.

By dynamic programming principle, G satisfies the following HJB equation:

$$-\sup_{\pi \in A} \mathcal{L}w(t, x, v) = 0 \tag{C.4}$$

for $(t, x, v) \in [0, T) \times (0, \infty)^2$ with terminal condition $w(T, x, v) = U(x)$, where $\mathcal{L}$ is the infinitesimal generator of processes X, V with control π, given by

$$\begin{aligned}
\mathcal{L}w(t, x, v) &= \frac{\partial w}{\partial t} + (r + (\mu - r)\pi)\, xw_x + \kappa(\theta - v)w_v \\
&\quad + \frac{1}{2}\pi^2 x^2 v w_{xx} + \frac{1}{2}\sigma^2 v w_{vv} + \rho\pi\sigma xv w_{xv}.
\end{aligned}$$

Remark C.0.4. Kraft (2005) shows that under the assumptions that (1) $\mu = r + \lambda V_t$ with λ constant, (2) $U(x) = \frac{x^\gamma}{\gamma}$ with $\gamma \in (0, 1)$ and (3) $\frac{\gamma}{1-\gamma}\lambda\left(\frac{\kappa\rho}{\sigma} + \frac{\lambda}{2}\right) < \frac{\kappa^2}{2\sigma^2}$, the

HJB equation (C.4) can be solved explicitly with value function given by

$$G(t, x, v) = \frac{x^\gamma}{\gamma} f(t, v)^c, \qquad (C.5)$$

where

$$f(t, v) = \exp\left\{\frac{\gamma}{c} r(T - t) - A(t, T) - B(t, T)v\right\},$$

and the optimal strategy given by

$$\pi^*(t) = \frac{1}{1 - \gamma}\lambda + \frac{\gamma}{(1 - \gamma)^2}\rho\sigma\lambda^2 \frac{e^{\tilde{a}(T-t)} - 1}{e^{\tilde{a}(T-t)}(\tilde{\kappa} + \tilde{a}) - \tilde{\kappa} + \tilde{a}}. \qquad (C.6)$$

Note that $A(t, T), B(t, T)$ are deterministic functions given by

$$A(t, T) = -\frac{\kappa\theta(\tilde{\kappa} - \tilde{a})}{\sigma^2}(T - t) + \frac{2\kappa\theta}{\sigma^2}\ln\left(\frac{1 - ke^{-\tilde{a}(T-t)}}{1 - k}\right),$$

$$B(t, T) = 2\tilde{\beta}\frac{e^{\tilde{a}(T-t)} - 1}{e^{\tilde{a}(T-t)}(\tilde{\kappa} + \tilde{a}) - \tilde{\kappa} + \tilde{a}},$$

and $c = \frac{1-\gamma}{1-\gamma+\rho^2\gamma}$, $\tilde{\beta} = -\frac{1}{2c}\frac{\gamma}{1-\gamma}\lambda^2$, $\tilde{a} = \sqrt{\tilde{\kappa}^2 + 2\tilde{\beta}\sigma^2}$, $\tilde{\kappa} = \kappa - \frac{\gamma}{1-\gamma}\rho\lambda\sigma$, $k = \frac{\tilde{\kappa}-\tilde{a}}{\tilde{\kappa}+\tilde{a}}$.

By Remark C.0.4, we have the theoretical result for the value function and optimal control under certain assumptions. Although constrained by the restricted assumptions, the theoretical result from Kraft (2005) can be used as benchmark, to test the convergence of our numerical result in the following sections.

Numerical Method Description

In the following sections, we will describe our numerical method based on the non-combining tree model.

Construction of the Approximating Tree

We construct the approximating tree model for the controlled process (X_t, V_t) following the virtue of Euler scheme in this section. We fix a time horizon $[t, T]$ where

$T > t$ is maturity time and a time discretization

$$\Delta t := \frac{T - t}{N},$$

where $N \geq 1$ is an integer. In the following sections, we assume $t = 0$ for simplicity.

Consider the discrete processes $\{S_k^N, V_k^N\}_{k=0}^N$ of the form

$$S_k^N := S_{k-1}^N \left(1 + \mu \Delta t + \sqrt{V_k^N \Delta t} \xi_k^S \right), \quad S_0^N = s, \tag{C.7}$$

and

$$V_k^N := V_{k-1}^N + \kappa(\theta - V_{k-1}^N)\Delta t + \sigma \sqrt{V_{k-1}^N \Delta t} \xi_k^V, \quad V_0^N = v, \tag{C.8}$$

where $\xi_k^V = \rho \xi_k^S + \sqrt{1 - \rho^2} \xi_k^Z$ and (ξ^S, ξ^Z)'s are random variables taking values in $\{-1, 1\}$ with probabilities $\frac{1}{2}$ respectively. For completeness, we assume that $\xi_0^S = \xi_0^Z = 0$.

Denote the value function under a N periods tree model setting by $G^N(t, x, v)$ and the wealth process by X^N, then we have

$$G^N(t, x, v) = \sup_{\pi \in \mathcal{A}^N} \mathbb{E}\left[U(X_T^{N,\pi}) \mid X_t^N = x, V_t^N = v \right], \tag{C.9}$$

where $\mathcal{A}^N$ is the admissible control set from time $t = 0$ to $T = N\Delta t$. The evolution of wealth process from $k\Delta t$ to $(k+1)\Delta t$ can be deduced from

$$\frac{X_{(k+1)\Delta t}^N - X_{k\Delta t}^N}{X_{k\Delta t}^N} = \pi_k \frac{S_{k+1}^N - S_k^N}{S_k^N} + r(1 - \pi_k)\Delta t.$$

Thus

$$X_{(k+1)\Delta t}^N = X_{k\Delta t}^N \left(1 + r\Delta t + \pi_k \left((\mu - r)\Delta t + \sqrt{V_k^N \Delta t} \xi_{k+1}^S \right) \right), \tag{C.10}$$

where π_k is the constant fraction of wealth into risky asset S from time step $k\Delta t$ to $(k+1)\Delta t$.

We will introduce the method to compute G^N which is an approxy of the real value function G defined by (C.3). The discrete optimal control at each step is a byproduct.

Computation Method for the Approximated Value Function

We numerically solve the approximated value function (C.9) by applying discrete dynamic programming principle. Define $\Theta := (\Theta_k)_{k=0}^N$ where $\Theta_k := (X_{k\Delta t}^N, V_k^N)$. Then the evolution from Θ_k to Θ_{k+1} is

$$
\Theta_{k+1} \mid \pi_k :=
\begin{cases}
\Theta_{k+1}^1(\pi_k) & \text{with prob. } 1/4, \\
\Theta_{k+1}^2(\pi_k) & \text{with prob. } 1/4, \\
\Theta_{k+1}^3(\pi_k) & \text{with prob. } 1/4, \\
\Theta_{k+1}^4(\pi_k) & \text{with prob. } 1/4,
\end{cases}
$$

where

$$
\Theta_{k+1}^1(\pi_k) = \left(X_{k\Delta t}^N \left(1 + r\Delta t + \pi_k \left((\mu - r)\Delta t + \sqrt{V_k^N \Delta t} \right) \right), V_{k+1}^N(1) \right),
$$

$$
\Theta_{k+1}^2(\pi_k) = \left(X_{k\Delta t}^N \left(1 + r\Delta t + \pi_k \left((\mu - r)\Delta t + \sqrt{V_k^N \Delta t} \right) \right), V_{k+1}^N(2) \right),
$$

$$
\Theta_{k+1}^3(\pi_k) = \left(X_{k\Delta t}^N \left(1 + r\Delta t + \pi_k \left((\mu - r)\Delta t - \sqrt{V_k^N \Delta t} \right) \right), V_{k+1}^N(3) \right),
$$

$$
\Theta_{k+1}^4(\pi_k) = \left(X_{k\Delta t}^N \left(1 + r\Delta t + \pi_k \left((\mu - r)\Delta t - \sqrt{V_k^N \Delta t} \right) \right), V_{k+1}^N(4) \right),
$$

and

$$
V_{k+1}^N(1) = V_k^N + \kappa(\theta - V_k^N)\Delta t + \sigma \sqrt{V_k^N \Delta t}(\rho + \sqrt{1 - \rho^2}),
$$

$$
V_{k+1}^N(2) = V_k^N + \kappa(\theta - V_k^N)\Delta t + \sigma \sqrt{V_k^N \Delta t}(\rho - \sqrt{1 - \rho^2}),
$$

$$V_{k+1}^N(3) = V_k^N + \kappa(\theta - V_k^N)\Delta t + \sigma\sqrt{V_k^N \Delta t}(-\rho + \sqrt{1 - \rho^2}),$$

$$V_{k+1}^N(4) = V_k^N + \kappa(\theta - V_k^N)\Delta t + \sigma\sqrt{V_k^N \Delta t}(-\rho - \sqrt{1 - \rho^2}).$$

Remark C.0.5. In implementation of the numerical method, we require the candidate control $\pi_k \in \mathbb{R}$ such that the wealth process is non-negative, which means

$$1 + r\Delta t + \pi_k\left((\mu - r)\Delta t - \sqrt{V_k^N \Delta t}\right) > 0.$$

Also note that there is chance that $V_{k+1}^N < 0$ in implementation. We set zero as the lower bound by standard truncation argument.

By dynamic programming principle (DPP), we have

$$G^N(k\Delta t, \Theta_k) = \sup_{\pi \in \mathcal{A}_k^N} \mathbb{E}\left[G^N\left((k+1)\Delta t, \Theta_{k+1}\right) \mid \Theta_k\right],$$

and

$$G^N\left((N-1)\Delta t, \Theta_{N-1}\right) = \sup_{\pi \in \mathcal{A}_{N-1}^N} \mathbb{E}\left[U(X_{N\Delta t}^{N,\pi}) \mid \Theta_{N-1}\right],$$

where $\mathcal{A}_k^N$ is the admissible control set from time $k\Delta t$ to $(k+1)\Delta t$.

Therefore, the numerical method to solve the approximated value function (C.9) can be summarized as the following steps:

- Step 1: $k = N - 1$.

$$
\begin{aligned}
&G^N\left((N-1)\Delta t, \Theta_{N-1}\right) \\
&= \sup_{\pi \in \mathcal{A}_{N-1}^N} \mathbb{E}\left[U(X_{N\Delta t}^{N,\pi}) \mid \Theta_{N-1}\right] \\
&= \frac{1}{2}\Bigg[U\left(X_{(N-1)\Delta t}^N\left(1 + r\Delta t + \left((\mu - r)\Delta t + \sqrt{V_{N-1}^N \Delta t}\right)\pi_{N-1}^*\right)\right) \\
&\qquad + U\left(X_{(N-1)\Delta t}^N\left(1 + r\Delta t + \left((\mu - r)\Delta t - \sqrt{V_{N-1}^N \Delta t}\right)\pi_{N-1}^*\right)\right)\Bigg],
\end{aligned}
$$

where

$$
\begin{aligned}
\pi^*_{N-1} \quad = \quad & \underset{\pi \in \mathcal{A}^N_{N-1}}{\operatorname{argmax}} \left[U \left(X^N_{(N-1)\Delta t} \left(1 + r\Delta t + \left((\mu - r)\Delta t + \sqrt{V^N_{N-1}\Delta t} \right) \pi \right) \right) \right. \\
& \left. + U \left(X^N_{(N-1)\Delta t} \left(1 + r\Delta t + \left((\mu - r)\Delta t - \sqrt{V^N_{N-1}\Delta t} \right) \pi \right) \right) \right].
\end{aligned}
$$

- Step 2: $0 \leq k < N - 1$.

$$
\begin{aligned}
& G^N \left(k\Delta t, \Theta_k \right) \\
= \quad & \sup_{\pi \in \mathcal{A}^N_k} \mathbb{E} \left[G^N \left((k+1)\Delta t, \Theta_{k+1} \right) \mid \Theta_k \right] \\
= \quad & \frac{1}{4} \left[G^N \left((k+1)\Delta t, \Theta^1_{k+1}(\pi^*_k) \right) + G^N \left((k+1)\Delta t, \Theta^2_{k+1}(\pi^*_k) \right) \right. \\
& \left. + G^N \left((k+1)\Delta t, \Theta^3_{k+1}(\pi^*_k) \right) + G^N \left((k+1)\Delta t, \Theta^4_{k+1}(\pi^*_k) \right) \right],
\end{aligned}
$$

where

$$
\begin{aligned}
\pi^*_k \quad = \quad & \underset{\pi \in \mathcal{A}^N_k}{\operatorname{argmax}} \left[G^N \left((k+1)\Delta t, \Theta^1_{k+1}(\pi) \right) + G^N \left((k+1)\Delta t, \Theta^2_{k+1}(\pi) \right) \right. \\
& \left. + G^N \left((k+1)\Delta t, \Theta^3_{k+1}(\pi) \right) + G^N \left((k+1)\Delta t, \Theta^4_{k+1}(\pi) \right) \right].
\end{aligned}
$$

The tree of variance process V for a fixed number of periods N is known at time 0, however the tree size increases rapidly due to the non-recombining property. For a N periods tree model, the number of V states for each period is given by Table C.1.

Period	1	2	3	4	5	6	7	8	...	N
Size	4	16	64	256	1024	4096	16384	65536	...	4^N

Table C.1: Tree size of variance process V.

Further, the tree of wealth process X is not fixed due to the continuous control π in the place. To reduce the computation size and make the numerical method

applicable in practice, we introduce a heuristic interpolation scheme for the wealth process X and variance process V. We will then adjust the numerical method described above based on the introduced interpolation scheme.

For the interpolation of variance process V, we firstly estimate the range of V size for each period. Note that we can get the exact range of V size as the tree is fixed at inception time. However, for a large N, the accurate range calculation is time consuming. Therefore, we estimate the range of V size by the following scheme:

$$V_{k,1} \leq V_k^N \leq V_{k,2},$$
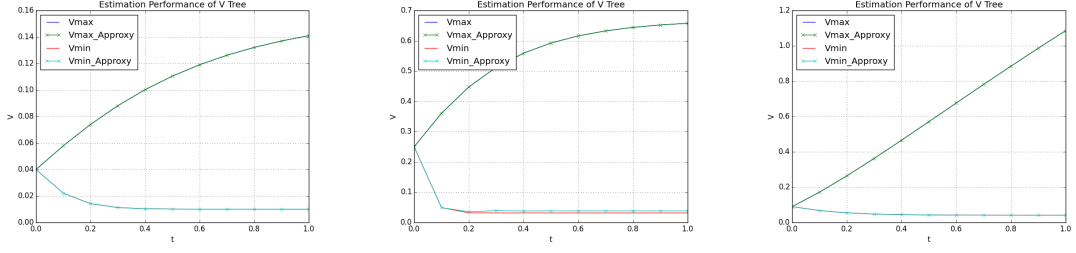
where $V_{0,1} = V_{0,2} = V_0$ and

$$V_{k,1} = V_{k-1,1} + \kappa(\theta - V_{k-1,1})\Delta t + \sigma\sqrt{V_{k-1,1}\Delta t}\left(-\rho \cdot \text{sign}(\rho) - \sqrt{1-\rho^2}\right),$$

$$V_{k,2} = V_{k-1,2} + \kappa(\theta - V_{k-1,2})\Delta t + \sigma\sqrt{V_{k-1,2}\Delta t}\left(\rho \cdot \text{sign}(\rho) + \sqrt{1-\rho^2}\right).$$

To check the estimation performance, we apply above estimation scheme to the following three sets of parameter values and compare with the corresponding accurate range for a 10 periods tree model.

- Parameter set 1: $\kappa = 3.0, \theta = 0.04, \sigma = 0.2, \rho = -0.7, V_0 = 0.04, T = 1, N = 10$;

- Parameter set 2: $\kappa = 5.0, \theta = 0.16, \sigma = 0.9, \rho = 0.1, V_0 = 0.25, T = 1, N = 10$;

- Parameter set 3: $\kappa = 1.15, \theta = 0.348, \sigma = 0.39, \rho = -0.64, V_0 = 0.09, T = 1, N = 10$.

Figure C.1 shows that the estimated range exactly coincides with the accurate range for parameter set 1 and 3. However, the lower range estimation is slightly above the true lower range for parameter set 2 which has a very large σ. Therefore, we fix our

Figure C.1: Comparison of V range estimation and accurate range

estimation of lower range such that the range estimation is more conservative:

$$V_{k,1} = 0.$$

In the following numerical test, we use the estimation range of V_k^N given by

$$V_{0,1} = V_{0,2} = V_0$$

and

$$\left[0, V_{k-1,2} + \kappa(\theta - V_{k-1,2})\Delta t + \sigma \sqrt{V_{k-1,2}\Delta t} \left(\rho \cdot \text{sign}(\rho) + \sqrt{1 - \rho^2} \right) \right].$$

for $k \geq 1$.

It is more tricky for the estimation of wealth process range as the wealth process depends on the control. Define π_{max} to be the largest control into consideration, then we can estimate the range of $X_{k\Delta t}$ for each time step $k\Delta t$. The range of $X_{k\Delta t}$ is estimated by

$$X_{k,1} \leq X_{k\Delta t}^N \leq X_{k,2},$$

where $X_{0,1} = X_{0,2} = X_0$ and

$$X_{k,1} = X_{k-1,1} \left(1 + r\Delta t + \left((\mu - r)\Delta t - \sqrt{V_{k-1,2}\Delta t} \right) \pi_{max} \right),$$

$$X_{k,2} = X_{k-1,2} \left(1 + r\Delta t + \left((\mu - r)\Delta t + \sqrt{V_{k-1,2}\Delta t} \right) \pi_{max} \right).$$

Based on the estimated range of X and V, we partition the range of $X_{k\Delta t}$ into M_k^X intervals with equal length such that the length of each small interval is approximately equal to $s_X > 0$, which means

$$M_k^X = \max\left\{\min\left\{\left\lceil\frac{X_{k,2} - X_{k,1}}{s_X}\right\rceil, M_{max}\right\}, M_{min}\right\},$$

where $\lceil x \rceil := \min\{n \in \mathbb{Z} \mid x \leq n\}$ and M_{max}, M_{min} are the maximum and minimum number of intervals. Intuitively, it means that the range of $X_{k\Delta t}$ is equally partitioned into several intervals with length approximately equal to s_X. The number of intervals does not exceed M_{max} for the benefit of computation speed and is not below M_{min} for the benefit of estimation accuracy. Similarly, we partition the range of V_k^N into M_k^V intervals where

$$M_k^V = \max\left\{\min\left\{\left\lceil\frac{V_{k,2} - V_{k,1}}{s_V}\right\rceil, M_{max}\right\}, M_{min}\right\}.$$

Denote the partition of $X_{k\Delta t}$ by $\mathbb{X}_k$ and the partition of V_k^N by $\mathbb{V}_k$.

Adapted Computation Method based on Interpolation Scheme

Based on the interpolation scheme, we adapt the numerical scheme into the following steps:

- Step 1: For each pair of $\Theta_{N-1} := \left(X_{(N-1)\Delta t}^N, V_{N-1}^N\right) \in \mathbb{X}_{N-1} \times \mathbb{V}_{N-1}$,

$$
\begin{aligned}
&G^N\left((N-1)\Delta t, \Theta_{N-1}\right) \\
&= \sup_{\pi \in \mathcal{A}_{N-1}^N} \mathbb{E}\left[U(X_{N\Delta t}^{N,\pi}) \mid \Theta_{N-1}\right] \\
&= \frac{1}{2}\left[U\left(X_{(N-1)\Delta t}^N\left(1 + r\Delta t + \left((\mu - r)\Delta t + \sqrt{V_{N-1}^N \Delta t}\right)\pi_{N-1}^*\right)\right)\right. \\
&\quad\left. + U\left(X_{(N-1)\Delta t}^N\left(1 + r\Delta t + \left((\mu - r)\Delta t - \sqrt{V_{N-1}^N \Delta t}\right)\pi_{N-1}^*\right)\right)\right],
\end{aligned}
$$

where

$$\begin{aligned}
\pi^*_{N-1} \;=\; & \operatorname*{argmax}_{\pi \in \mathcal{A}^N_{N-1}} \left[U\left(X^N_{(N-1)\Delta t}\left(1 + r\Delta t + \left((\mu - r)\Delta t + \sqrt{V^N_{N-1}\Delta t}\,\right)\pi\right)\right) \right. \\
& \left. + U\left(X^N_{(N-1)\Delta t}\left(1 + r\Delta t + \left((\mu - r)\Delta t - \sqrt{V^N_{N-1}\Delta t}\,\right)\pi\right)\right)\right].
\end{aligned}$$

Thus we get the discretized value functions $G^N\left((N-1)\Delta t, \Theta_{N-1}\right)$ for each $\Theta_{N-1} \in \mathbb{X}_{N-1} \times \mathbb{V}_{N-1}$. Applying the standard bilinear interpolation methodology to $G^N\left((N-1)\Delta t, \Theta_{N-1}\right)$, we get the interpolation function $\tilde{G}_{N-1}(\tilde{\Theta})$ for $\tilde{\Theta} \in [0,\infty) \times [0,\infty)$. For $\tilde{\Theta} := (x, v) \in (x_1, x_2) \times (v_1, v_2)$ where (x_1, v_1), (x_2, v_1), (x_1, v_2), $(x_2, v_2) \in \mathbb{X}_{N-1} \times \mathbb{V}_{N-1}$, we have

$$\tilde{G}_{N-1}(\tilde{\Theta}) := \frac{1}{(x_2 - x_1)(v_2 - v_1)}(x_2 - x, x - x_1)D\begin{pmatrix} v_2 - v \\ v - v_1 \end{pmatrix},$$

where

$$D := \begin{pmatrix} G^N\left((N-1)\Delta t, (x_1, v_1)\right) & G^N\left((N-1)\Delta t, (x_1, v_2)\right) \\ G^N\left((N-1)\Delta t, (x_2, v_1)\right) & G^N\left((N-1)\Delta t, (x_2, v_2)\right) \end{pmatrix}.$$

- Step 2: For each pair of $\Theta_k := \left(X^N_{k\Delta t}, V^N_k\right) \in \mathbb{X}_k \times \mathbb{V}_k$ where $0 \leq k < N-1$,

$$\begin{aligned}
& G^N\left(k\Delta t, \Theta_k\right) \\
=\; & \sup_{\pi \in \mathcal{A}^N_k} \mathbb{E}\left[G^N\left((k+1)\Delta t, \Theta_{k+1}\right) \mid \Theta_k\right] \\
=\; & \frac{1}{4}\left[G^N\left((k+1)\Delta t, \Theta^1_{k+1}(\pi^*_k)\right) + G^N\left((k+1)\Delta t, \Theta^2_{k+1}(\pi^*_k)\right) \right. \\
& \left. + G^N\left((k+1)\Delta t, \Theta^3_{k+1}(\pi^*_k)\right) + G^N\left((k+1)\Delta t, \Theta^4_{k+1}(\pi^*_k)\right)\right] \\
\approx\; & \frac{1}{4}\left[\tilde{G}_{k+1}(\Theta^1_{k+1}(\pi^*_k)) + \tilde{G}_{k+1}(\Theta^2_{k+1}(\pi^*_k)) \right. \\
& \left. + \tilde{G}_{k+1}(\Theta^3_{k+1}(\pi^*_k)) + \tilde{G}_{k+1}(\Theta^4_{k+1}(\pi^*_k))\right],
\end{aligned}$$

where

$$\pi_k^* \;=\; \underset{\pi \in \mathcal{A}_k^N}{\mathrm{argmax}} \left[\tilde{G}_{k+1}(\Theta_{k+1}^1(\pi)) + \tilde{G}_{k+1}(\Theta_{k+1}^2(\pi)) \right.$$

$$\left. + \tilde{G}_{k+1}(\Theta_{k+1}^3(\pi)) + \tilde{G}_{k+1}(\Theta_{k+1}^4(\pi)) \right].$$

Thus we get the discretized value functions $G\left(k\Delta t, \Theta_k\right)$ for each $\Theta_k \in \mathbb{X}_k \times \mathbb{V}_k$. Applying the same bilinear interpolation methodology (specified in Step 1) to the pair $(\Theta_k, G\left(k\Delta t, \Theta_k\right))$, then we get the interpolation function $\tilde{G}_k(\tilde{\Theta})$ for $\tilde{\Theta} \in [0, \infty) \times [0, \infty)$.

In the following sections, we investigate the convergence of our numerical method by comparing the numerical results with the theoretical results from Kraft (2005) under power utility, and the tight bounds from Ma et al. (2017) under non-HARA utility, respectively.

Convergence of Numerical Method for Power Utility

We compare the numerical results with power utility function with the closed form solution in Kraft (2005). We set $T = 1, X_0 = 1, \pi_{max} = 3, M_{min} = 3, M_{max} = 10000, s_X = 0.1, s_V = 0.001$ and the following 2 parameter value sets:

- Parameter set 1: $\lambda = 1.0, r = 0.05, \kappa = 3.0, \theta = 0.04, \sigma = 0.2, \rho = -0.7, V_0 = 0.04$.

- Parameter set 2: $\lambda = 0.5, r = 0.05, \kappa = 1.15, \theta = 0.348, \sigma = 0.39, \rho = -0.64, V_0 = 0.09$.

Note that the 2 parameter sets are calibrated from the real market data, see Akyildirim et al. (2014).

We compare the first time step control π_0^* and value function $G^N(0, \Theta_0)$ with the theoretical results, under different choices of number of periods N. Due to the

running time consumption issue, we choose the largest tested N to be 30. The comparison result of parameter set 1 is given by Table C.2.

N	1	2	3	4	5	8
π_0^*	2.1875	2.0216	2.0155	1.9990	1.9914	1.9792
G^N	2.09165	2.09013	2.09082	2.09074	2.09074	2.09076
N	10	15	20	25	30	Kraft
π_0^*	1.9744	1.9695	1.9654	1.9628	1.9589	1.9573
G^N	2.09077	2.09078	2.09079	2.09079	2.09079	2.09079

Table C.2: Numerical Results of Parameter set 1 with $M_{max} = 10000, (s_X, s_V) = (0.1, 0.001)$.

Table C.2 shows clearly that π_0^* and $G^N(0, \Theta_0)$ converge to Kraft's real solutions when N increases. Actually, the value function G^N converges faster than the optimal control. Table C.3 confirms the same convergence pattern of parameter set 2.

N	1	2	3	4	5	8
π_0^*	1.0742	1.0037	0.9900	0.9831	0.9797	0.9720
G^N	2.07284	2.09239	2.09516	2.09635	2.09700	2.09792
N	10	15	20	25	30	Kraft
π_0^*	0.9747	0.9713	0.9701	0.9667	0.9654	0.9646
G^N	2.09809	2.09821	2.09832	2.09928	2.09930	2.09930

Table C.3: Numerical Results of Parameter set 2 with $M_{max} = 10000, (s_X, s_V) = (0.1, 0.001)$.

Note that the interpolation scheme introduces inaccuracy into the numerical method. We further test the sensitivity of convergence speed to the roughness of the interpolation scheme. Intuitively the roughness of the interpolation scheme is determined by M_{max} and (s_X, s_V).

Table C.4 shows the numerical result with $M_{max} = 1000$. Compared with Table C.2, it is clear that although the approximated value function G^N still converges to the real value function G, the convergence speed is slower. This is due to the fact that with a smaller M_{max}, the interpolation grid size becomes larger when N is large enough.

Table C.5 shows the numerical result with $(s_X, s_V) = (0.2, 0.002)$.

N	1	2	3	4	5	8
π_0^*	2.1875	2.0218	2.0158	1.9994	1.9904	1.9777
G^N	2.09165	2.09013	2.09081	2.09074	2.09074	2.09076
N	10	15	20	25	30	Kraft
π_0^*	1.9750	1.9687	1.9661	1.9650	1.9639	1.9573
G^N	2.09077	2.09077	2.09078	2.09078	2.09079	2.09079

Table C.4: Numerical Results of Parameter set 1 with $M_{max} = 1000, (s_X, s_V) = (0.1, 0.001)$.

N	1	2	3	4	5	8
π_0^*	2.1875	2.0217	2.0139	1.9994	1.9923	1.9785
G^N	2.09165	2.09018	2.09051	2.09065	2.09070	2.09073
N	10	15	20	25	30	Kraft
π_0^*	1.9764	1.9722	1.9697	1.9680	1.9669	1.9573
G^N	2.09074	2.09075	2.09077	2.09078	2.09079	2.09079

Table C.5: Numerical Results of Parameter set 1 with $M_{max} = 10000, (s_X, s_V) = (0.2, 0.002)$.

The numerical comparison between the result from our numerical method and the theoretical result from Kraft (2005) shows that for power utility function, the approximated value function and optimal control from our numerical result seems to converge to the real solutions, and the convergence speed depends on the interpolation scheme. Precisely, the convergence speed increases with M_{max} while decreases with s_X and s_V.

Convergence of Numerical Method for Non-HARA Utility

We compare the numerical results with a non-HARA utility function. We choose the following non-HARA utility function given in Ma et al. (2017):

$$U(x) = \frac{1}{3}H(x)^{-3} + H(x)^{-1} + xH(x),$$

for $x > 0$, where

$$H(x) = \left(\frac{2}{-1 + \sqrt{1 + 4x}} \right)^{1/2}.$$

We firstly check the convergence of our numerical method with non-HARA utility when $\sigma = 0$. Note that under the assumption that $\sigma = 0$, the Heston model degenerates to Black-Scholes model and we can derive the closed-form solution under this extreme case. We set the following benchmark parameter values (from Ma et al. (2017)):

$$\lambda = 0.5, r = 0.05, \kappa = 10, \theta = 0.05, \sigma = 0, \rho = -0.5, V_0 = 0.05, T = 1, X_0 = 1.$$

Table C.6 shows the comparison of the approximated value function from our numerical method with the real solution.

N	1	2	3	4	5	8
G^N	2.30618	2.30698	2.30725	2.30739	2.30747	2.30761
N	10	15	20	25	30	G
G^N	2.30764	2.30770	2.30775	2.30779	2.30781	2.30781

Table C.6: Numerical Results of Non-HARA utility under Black-Scholes model with with $\pi_{max} = 3, M_{max} = 10000, (s_X, s_V) = (0.1, 0.001)$.

It is clear that the approximated value function converges to the real solution. Also note that in Ma et al. (2017), the estimated lower bound and upper bound are 2.30769 and 2.30784 respectively, thus the approximated value function lies in the estimated range when $N \geq 15$.

Next we test the real Heston model case where $\sigma \neq 0$, for which we don't have closed form solution from the literature. We compare the numerical result from our approach with the estimated bounds from Ma et al. (2017). We use the following benchmark parameter values (also from Ma et al. (2017)):

$$\lambda = 0.5, r = 0.05, \kappa = 10, \theta = 0.05, \sigma = 0.5, \rho = -0.5, V_0 = 0.5, T = 1, X_0 = 1.$$

Table C.7 shows the comparison results. It is clear that the approximated value function lies in the estimated range (LB: lower bound, UB: upper bound) when

$N \geq 20$.

N	1	2	3	4	5	8
G^N	2.53466	2.40249	2.36203	2.35347	2.34764	2.33018

N	10	15	20	25	30	LB	UB
G^N	2.33003	2.32946	2.32803	2.32775	2.32773	2.32757	2.32786

Table C.7: Numerical Results of Non-HARA utility under Heston model with with $\pi_{max} = 3, M_{max} = 10000, (s_X, s_V) = (0.1, 0.001)$.

The numerical comparison between the result from our numerical method and the estimated bounds from Ma et al. (2017) shows that for non-HARA utility function, the approximated value function from our numerical result seems to converge to the real value function.

Conclusions

In this appendix, we introduce a numerical method to solve a dynamic portfolio optimization problem under Heston model framework. We approximate the controlled process by a non-combining tree model and apply discrete dynamic programming principle to the tree model directly. We numerically show the convergence of our approach, by comparing the approximated value function and optimal control with the theoretical solutions from Kraft (2005) under power utility and the tight estimated bounds from Ma et al. (2017). We also show that the convergence performance is sensitive to the interpolation scheme. There remain many open questions in this numerical approach. For example, the convergence result reported in this appendix is only based on numerical investigation. The theoretical proof of convergence should be considered in the future research.

D

Notations and Theorems

D.1 General Notations and Abbreviations

For any real numbers x, y,

$$x^+ = \max(x, 0), \quad x^- = \max(-x, 0), \quad x \wedge y = \min(x, y), \quad x \vee y = \max(x, y).$$

$\mathbb{R}^d$ denotes the d–dimensional Euclidian space. $\mathbb{R} = \mathbb{R}^1$. For all $x = (x_1, \ldots, x_d), y = (y_1, \ldots, y_d)$ in $\mathbb{R}^d$, we denote by $\langle \cdot, \cdot \rangle$ the inner product and by $|| \cdot ||$ the Euclidian norm:

$$\langle x, y \rangle = \sum_{i=1}^{d} x_i y_i, \quad ||x|| = \sqrt{\langle x, x \rangle}.$$

$\mathcal{S}^{n \times d}$ is the set of real-valued $n \times d$ matrices. I_n is the $n \times n$ identity matrix. For all $\Sigma = (\sigma^{ij})_{1 \leq i \leq n, 1 \leq j \leq d} \in \mathcal{S}^{n \times d}$, we denote by $\Sigma^T = (\sigma^{ji})_{1 \leq j \leq d, 1 \leq i \leq n}$ the transpose matrix in $\mathcal{S}^{d \times n}$. We set $\text{Tr}(\Sigma) = \sum_{i=1}^{n} \sigma^{ii}$ the trace of a $n \times n$ matrix $\Sigma \in \mathcal{S}^{n \times n}$. We

say that $\Sigma_1 \leq \Sigma_2$ for $\Sigma_1, \Sigma_2 \in \mathcal{S}^{n \times n}$ if $\Sigma_2 - \Sigma_1$ is nonnegative definite.

We denote by $B_\eta(x)$ (resp. $\bar{B}_\eta(x)$) the open (resp. closed) ball of center $x \in \mathbb{R}^d$, and radius $\eta > 0$.

$C([0, T] \times \mathcal{O})$ is the space of all real-valued continuous functions on $[0, T] \times \mathcal{O}$.
$C^{1,2,\dots,2}([0, T) \times \mathcal{O})$ is the space of all real-valued continuous functions f on $[0, T) \times \mathcal{O}$ whose partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x_i}$, $\frac{\partial^2 f}{\partial x_i \partial x_j}$ exist and are continuous on $[0, T) \times \mathcal{O}$.
$f(x) = o(g(x))$ means that $\lim_{x \to 0} f(x)/g(x) = 0$.

$(\Omega, \mathcal{F}, \mathbb{P})$: probability space.
$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$: filtered probability space.
$\mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t]$: expectation of random variable X under measure $\mathbb{Q}$ given filtration $\mathcal{F}_t$.
$\frac{d\mathbb{Q}}{d\mathbb{P}}$: Radon-Nikodym density.

SDE: stochastic differential equation

BSDE: backward stochastic differential equation

PDE: partial differential equation

DPP: dynamic programming principle

HJB: Hamilton-Jacobi-Bellman

CoCo: contingent convertible

CET1: common equity tier 1

PONV: point of non-viability

RWA: risk-weighted asset

HARA: hyperbolic absolute risk aversion

u.s.c.: upper-semicontinuous

l.s.c.: lower-semicontinuous

D.2 DEFINITIONS AND THEOREMS

The utility function is used to represent investor's risk averse level. It is usually assumed that the utility function $U(\cdot) \in C^2$ has the following properties: strictly concave and increasing. Under these assumptions, the relation between risk averse and utility function defined by

Definition D.2.1. (Absolute risk aversion) The Arrow-Pratt measure of absolute risk-aversion is defined as

$$A(x) := -\frac{U''(x)}{U'(x)}.$$

Based on Definition D.2.1, the Arrow-Pratt measure of log utility $U(x) = \ln x$ is $1/x$ and that of power utility $U(x) = \frac{x^\gamma}{\gamma}$ is $(1-\gamma)/x$.

Definition D.2.2. (HARA utility function) A function $U(\cdot)$ is said to be a HARA (hyperbolic absolute risk aversion) function if it admits the representation

$$U(x) = \frac{1-\gamma}{\gamma} \left(\frac{\beta}{1-\gamma} x + \eta \right)^\gamma,$$

where $\gamma < 1, \gamma \neq 0, \beta > 0, \frac{\beta}{1-\gamma} x + \eta > 0$.

The following generalized Ito's formula is cited from Protter (2005) and it is used throughout this thesis.

Theorem D.2.3. *(Generalized Ito's formula) Let X be a semimartingale and let f be a C^2 real function. Then $f(X)$ is again a semimartingale, and the following formula holds:*

$$f(X_t) = f(X_0) + \int_0^t f'(X_{s-}) dX_s^C + \frac{1}{2} \int_0^t f''(X_{s-}) d[X,X]_s^C + \sum_{0 \leq s \leq t} \Delta f(X_s),$$

where X^C is the continuous part of process X and $\Delta f(X_s) := f(X_s) - f(X_{s-})$.

The following existence and uniqueness result of one-dimensional SDE strong solution is used as the foundation of the wealth process definition throughout this thesis. The theorem is cited from Øksendal (2003).

Theorem D.2.4. *(Existence and uniqueness of SDE strong solution) Let us suppose that the coefficients of the one-dimensional SDE*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \tag{D.1}$$

satisfies the conditions

$$|b(t, x) - b(t, y)| \leq K|x - y|,$$

$$|\sigma(t, x) - \sigma(t, y)| \leq h(|x - y|),$$

for every $0 \leq t < \infty$ and $x, y \in \mathbb{R}$, where K is a positive constant and $h : [0, \infty) \to [0, \infty)$ is a strictly increasing function with $h(0) = 0$ and

$$\int_0^\epsilon h^{-2}(u)du = \infty, \quad \forall \epsilon > 0.$$

Then there exists a unique strong solution to (D.1).

The following Doob's sub-martingale inequality is used in the proof of Theorem 2.4.2 in Chapter 1. The theorem is cited from Pham (2009).

Theorem D.2.5. *(Doob's sub-martingale inequality) Let X be a sub-martingale that takes nonnegative real values, then for a constant $K > 0$,*

$$\mathbb{P}\left[\sup_{0 \leq t \leq T} X_t \geq K\right] \leq \frac{\mathbb{E}[X_T]}{K},$$

and consequently

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |X_t|^p\right] \leq \left(\frac{p}{1-p}\right)^p \mathbb{E}\left[|X_T|^p\right].$$

The following Crandall-Ishii's lemma is used in the proof of comparison principle (Theorem 2.8.5). The Lemma is cited from Crandall et al. (1992).

Lemma D.2.6. *(Crandall-Ishii's lemma) Let U (resp. V) be a u.s.c. (resp. l.s.c.) function on $[0,T) \times \mathbb{R}^n$, $\phi \in C^{1,1,2,2}([0,T)^2 \times \mathbb{R}^n \times \mathbb{R}^n)$, and $(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \in [0,T)^2 \times \mathbb{R}^n \times \mathbb{R}^n$ a local maximum of $U(t,x) - V(t,y) - \phi(t,s,x,y)$. Then, for all $\eta > 0$, there exist $M, N \in \mathcal{S}^{n \times n}$ satisfying*

$$\left(\frac{\partial \phi}{\partial t}(\bar{t}, \bar{s}, \bar{x}, \bar{y}), D_x \phi(\bar{t}, \bar{s}, \bar{x}, \bar{y}), M \right) \in \bar{\mathcal{P}}^{2,+} U(\bar{t}, \bar{x}),$$

$$\left(-\frac{\partial \phi}{\partial s}(\bar{t}, \bar{s}, \bar{x}, \bar{y}), -D_y \phi(\bar{t}, \bar{s}, \bar{x}, \bar{y}), N \right) \in \bar{\mathcal{P}}^{2,-} V(\bar{s}, \bar{y}),$$

and

$$\begin{pmatrix} M & 0 \\ 0 & -N \end{pmatrix} \leq D_{x,y}^2 \phi(\bar{t}, \bar{s}, \bar{x}, \bar{y}) + \eta \left(D_{x,y}^2 \phi(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \right)^2,$$

where $\bar{\mathcal{P}}^{2,+} U(\bar{t}, \bar{x})$ (resp. $\bar{\mathcal{P}}^{2,-} V(\bar{s}, \bar{y})$) are the limiting second-order superjet (resp. subjet) of U (resp. V).

Remark D.2.7. By choosing $\eta = \epsilon$ and $\phi(t,s,x,y) = \frac{1}{2\epsilon}\left(|t-s|^2 + |x-y|^2\right)$, we obtain

$$\begin{pmatrix} M & 0 \\ 0 & -N \end{pmatrix} \leq \frac{1}{\epsilon} \begin{pmatrix} I_n & -I_n \\ -I_n & I_n \end{pmatrix}.$$

The following Martingale representation theorem is cited from Protter (2005). It is used in the derivation of CoCo bond dynamic under risk-neutral measure.

Theorem D.2.8. *(Martingale representation theorem) Let $(W_t)_{t \geq 0}$ be a Brownian motion on a standard filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and let $\bar{\mathcal{F}}_t$ be the augmented filtration generated by W. If X_T is a square integrable random variable measurable with respect to $\bar{\mathcal{F}}_T$, then there exists a predictable process λ_t which is adapted with respect to $\bar{\mathcal{F}}_t$, such that*

$$X_T = \mathbb{E}[X_T] + \int_0^T \lambda_t dW_t.$$