# FINITE HORIZON PORTFOLIO SELECTION WITH TRANSACTION COSTS

LI PEIFAN

NATIONAL UNIVERSITY OF SINGAPORE

2009

# FINITE HORIZON PORTFOLIO SELECTION WITH TRANSACTION COSTS

**LI PEIFAN**

*(M.Sci., Peking University)*

**A THESIS SUBMITTED**

**FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**DEPARTMENT OF MATHEMATICS**

**NATIONAL UNIVERSITY OF SINGAPORE**

**2009**

# Acknowledgements

The last four years have been one of the most important stages in my life. The experience in my Ph.D. period will benefit me for a lifetime. I would like to take this opportunity to express my immense gratitude to all those who have kindly helped me and all those who have made my graduate life at NUS both productive and enjoyable.

At the very first, I am honored to express my deepest gratitude to my dedicated supervisor, A. Prof. DAI Min. This thesis would not have been possible without his able supervision. He has offered me a great many of invaluable ideas and great suggestions with his insightful discoveries, profound knowledge, and rich research experience. From him, I learn not only the knowledge, but also the professional ethics, both of which will stay with me for many years to come. His encouragement, patience and kindness through all these years are greatly appreciated and I am very much obliged to his efforts of helping me finish the dissertation.

This thesis mainly contains two parts, each of which is from a research paper. I am deeply indebted to the co-authors. Besides my advisor A. Prof. DAI Min, they are Professor JIANG Lishang from Tongji University and Professor YI Fahuai from South China Normal University for the first paper on optimal portfolio selection with consumption, and Professor LIU Hong from Washington University in St. Louis for the second paper on liquidity premium under market closure. I owe special thanks to Professor LIU Hong for offering me numerous great ideas to complete this thesis. His insight and wide eyeshot deeply impressed me, as well as his meticulous attitude in research. I would also like to thank Dr. JIN Hanqing from helpful discussion and insightful suggestions.

My great gratitude also goes to Mr. ZHONG Yifei and Dr. YANG Zhou, who have been selflessly and generously sharing their insights and ideas with me. Their kindness are always appreciated. I would also like to thank Mr. WANG Shengyuan for proofreading the thesis.

I have many thanks to my fellow postgraduate friends, who shared the experience at NUS with me. Thanks for accompanying me these years, for making my graduate life joyful, and for always being there when needed.

Last, but certainly not the least, I would like to thank my family. I want to express my gratitude to my dearest husband, for his unceasing love and continuous support. I also want to thank my parents for their love and support all the way.

**Li Peifan**

**August 2009**

# Contents

# Summary

This thesis concerns continuous-time portfolios selection for a constant relative risk aversion (CRRA) investor who faces proportional transaction costs and a finite time horizon. Mathematically, it is a singular stochastic control problem whose value function satisfies a parabolic variational inequality with gradient constraints. The problem gives rise to two free boundaries which stand for the optimal buying and selling strategies, respectively. Two factors are considered separately in this thesis: consumption and market closure. In the consumption case, we present an analytical approach to analyze the behaviors of the free boundaries. The regularity of the value function is studied as well. In the market closure case, we find that assuming the well-established time-varying return dynamics can generate a first order effect of transaction costs on liquidity premium, which is much greater than that found by existing literature and comparable to empirical evidence. The impacts of market closure on trading strategies, wealth loss, and trading volume are investigated in details.

# List of Tables

# List of Figures

Chapter 1

# Introduction

People make investments to accumulate wealth. In the booming financial markets, people are offered various choices for investments other than the traditional way of depositing money in banks. To maximize their utilities from investment, investors face the problem of setting up portfolios and selecting the components among risk-free assets (e.g. government bonds and deposits) and risky ones (e.g. stocks). Generally speaking, a risk-free asset guarantees some deterministic return rate for investors; while a risky asset provides a stochastic return rate, whose expected value is usually higher than the risk-free return rate. However, due to the random nature, the realized return rate of a risky asset deviates from its expected value almost all the time, so investors may not always get higher return from a stock than the risk-free rate. This uncertainty necessitates tradeoff between risk-free and risky assets in a portfolio.

## 1.1 Review on portfolio selection with transaction costs

The portfolio selection problem has received extensive attention from researchers. The key words "portfolio selection" hit as many as 27,800 records on "Google Scholar". However, the methods used by the articles are quite similar. Most studies deal with portfolio selection problem with two approaches: Mean-Variance Optimization, and direct utility maximization. Harry M. Markowitz, who was awarded the 1990 Nobel Prize in Economics for his pioneering work in modern portfolio selection theory, initiated the Mean-Variance Optimization approach (Markowitz 1952, Markowitz 1956, Markowitz 1959). This approach established a tradeoff between reward and risk by maximizing an investor's expected return subject to a selected level of risk. Later on, scholars realized that essentially, the Mean-Variance Optimization approach was highly related to utility maximization. In many circumstances, to implement the Mean-Variance Optimization was equivalent to conduct some special utility maximization (Kroll, et. al. 1984). Merton (1971) first formulated the portfolio selection problem in the framework of utility maximization. He showed that for an investor with constant relative risk aversion (CRRA) utility function, the optimal trading strategy was to keep a constant fraction of total wealth in stock. However, this work was based on the assumption that no transaction costs applied and that the investment horizon was infinite. In reality, transaction costs do exist, and Merton's strategy is impractical because of the innumerous cost generated by incessant trading.

To overcome the impracticability of Merton's strategy, Magil and Constantinides (1976) took into account proportional transaction cost, the amount to be paid upon transaction which is proportional to the value of stock purchased or sold. They proposed that with the presence of transaction cost, an investor should never trade if the fraction of stock in total wealth is kept within a range. They also suggested that an investor should sell some stock if the stock fraction exceeded the upper bound, or purchase some stock if the stock fraction dropped below the lower bound. They name this range of stock fraction as the "no-trading region" or "no-transaction region", and defined its upper or lower bound as the selling or buying boundary, respectively.

It was Davis and Norman (1990) that first formulated the problem of portfolio selection with transaction costs as a free boundary problem, where the boundary of the no-trading region was the so-called free boundary. They then studied the properties of the free boundary that reflected the optimal strategy.

Shreve and Soner (1994) came next. In terms of a viscosity solution approach, they entirely characterized the behaviors of the free boundary. Akian, Menaldi, and Sulem (1996) considered an extension to the case of multiple risky assets. Janeček and Shreve (2004) presented an asymptotic expansion of the associated value function and obtained some asymptotic results on the free boundary. All of these works were confined to infinite horizon problems. Besides the above mentioned ones, there are still enormous papers on transaction costs, including Shreve, Soner and Xu (1991), Bertsimas and Lo (1998), Liu (2004), Cocco (2005), and Gomes and Michaelides (2005).

It has been challenging to take the finite horizon case into consideration since the corresponding free boundary (optimal trading strategy) would vary with time. Theoretical analysis on the finite horizon problem became possible only very recently. For example, Liu and Loewenstein (2002) examined the optimal strategy by virtue of a sequence of analytical solutions that converged to the solution of the finite horizon optimal investment problem with transaction costs. Dai and Yi (2009) considered the same problem and derived an equivalent variational inequality, by which they completely figured out the optimal strategy. Dai, Xu, and Zhou (2008) extended the idea of Dai and Yi (2009) to the continuous-time mean-variance analysis with transaction costs. By bootstrap technique, they proved infinite smoothness of the free boundary.

It is rather challenging to incorporate consumption to the finite horizon portfolio selection. Dai, Jiang and Yi (2007) tried to employ the methodology in Dai and Yi (2009) for investigation on the impact of consumption. They presented a complete analysis on the regularity of solution and the behaviors of free boundaries. However, their approach was based on some technical condition, which would not always be reasonable.

## 1.2 Review on liquidity premium

### 1.2.1 Equity premium puzzle

Generally speaking, investors are risk averse, therefore they demand compensation for holding risky assets. That is why the expected return on equities, which are volatile, is usually greater than the return on bonds, which are risk-free. Economists define the difference between risky and risk-free return as equity premium. One interesting phenomenon about equity premium is that the real equity premium observed in financial industry is significantly higher than the theoretical equity premium calculated from existing economics models. Mehra and Prescott (1985) first observed this phenomenon by analyzing the 1889-1978 S&P 500 Indexes. They found that the average equity premium was 6.18%, while the highest premium that could be calculated from economic models was 0.35%, which was significantly lower. Mehra and Prescott's discovery on equity premium was reconfirmed by most following research. Economists named this phenomenon as the "equity premium puzzle".

### 1.2.2 Liquidity premium

As a consequence of paying out transaction costs at trading, investors would expect compensation from equities' return. Liquidity premium, defined as the return compensation due to transaction cost, should contribute to equity premium. While economists are seeking rational solutions diligently for the equity premium puzzle, liquidity premium seems to be a highly likely answer.

However, most portfolio selection models (e.g., Constantinides (1986)) concluded that the liquidity premium (i.e., the maximum expected return an investor is willing to exchange for zero transaction cost) was an order of magnitude smaller than transaction cost. For example, Constantinides (1986) found that the liquidity premium to transaction cost (LPTC) ratio was only about 0.14 with a proportional transaction cost of 1%. The main intuition behind this conclusion was that with constant return dynamics, investors did not need to trade often and thus the loss from paying transaction costs was small.

However, this finding sharply contrasted with many empirical studies that suggested the importance of transaction costs or related measures such as turnover in influencing the cross-sectional patterns of expected returns. For example, Amihud and Mendelson (1986) found that the LPTC ratio was about 2.4, while Eleswarapu (1997) found it was about 0.9. Assuming that return dynamics varied across bull and bear economic regimes, Jang et. al. (2007) showed that transaction costs can have a significantly larger effect on liquidity premia because of the necessity to trade more frequently. However, Jang et. al. (2007) still assumed that in a given regime, return dynamics remained the same across trading and nontrading periods. Since bull and bear regimes switched infrequently and volatilities across these regimes did not differ too much, the liquidity premium found by Jang et. al. (2007) given reasonable calibration was about 0.5, which was still small relative to that suggested by empirical evidence.

### 1.2.3 Market closure and time-varying return dynamics

As we go through the literature, we find that most of the existing portfolio selection models assume that market is continuously open and stock return dynamics is constant across trading and nontrading periods. (e.g. Merton (1987), Constantinides (1986), Vayanos (1998), Liu and Loewenstein (2002), and Liu (2004).) One of the important implications of this assumption is that transaction costs only have a second-order effect for asset pricing.

However, market closures during nights, weekends, and holidays are implemented in almost all financial markets. With this periodic opening and closing of market, the return dynamics of stock would like to also change periodically in time. An extensive literature on stock return dynamics across trading and nontrading periods found that while expected returns did not vary significantly across these periods, volatilities did. For example, French and Roll (1986) and Stoll and Whaley (1990) found that volatility during trading periods was more than four times the volatility during non-trading periods on a per-hour basis. Furthermore, French and Roll (1986) found that the principle factor behind high trading-time variances was the private information revealed by informed trades during trading hours, although mispricing also contributed to it.

## 1.3 Purpose and scope of this thesis

We are going to study two problems related to finite horizon portfolio selection in this thesis. First comes the investigation of consumption and next is the study on liquidity premium with market closure.

Dai and Yi (2009) solved the problem of finite horizon portfolio selection with transaction costs in the absence of consumption. To our best knowledge, that is the only paper that directly investigated the finite horizon portfolio selection problem by present.[1] However, with consumption involved, the model would become more complicated. Following the approach in Dai and Yi (2009), Dai, Jiang and Yi (2007) tried to investigate the impact of consumption on optimal investment. As mentioned before, they obtained the regularities of solution and figured out the behaviors of free boundaries. However, their arguments followed Friedman (1975), so to prove infinite smoothness of the free boundary, one must ensure monotonicity in time of the value function (i.e. $v_t \geq 0$). Furthermore, to prove $v_t \geq 0$, they imposed some technique conditions, which would not always be reasonable. In Dai and Zhong (2009), they showed that the value function was NOT always monotone in time without these technical conditions.

As an extension of Dai, Jiang and Yi (2007), this thesis aims at establishing proof for smoothness of free boundary without the technical conditions. We will follow the idea in Dai, Xu and Zhou (2008) to prove cone property of the value function, which further leads to smoothness. In this way, we avoid relying on the monotonicity in time (which might be unavailable in some cases) any more. And the smoothness of free boundary consolidates other arguments for the optimal investment problem with consumption. This part of the thesis is based on Dai et. al. (2009) with the following objectives:

- To investigate the optimal investment problem with consumption with PDE methods;

[1] Liu and Loewenstein (2002) also examined the finite horizon problem by an indirect approach. They took a sequence of analytical solutions that converged to the solution to the finite horizon optimal investment problem with transaction costs.

- To prove the regularity of value functions;
- To characterize the optimal trading strategies.

Regarding liquidity premium, the effects of open-close mechanism in financial markets and the time-varying return dynamics across trading and non-trading periods have remained unclear. In this thesis, we will also try to bridge the gaps by introducing two factors to the finite horizon optimal investment model: a) market closure, and b) time-varying stock return dynamics (dynamic opportunity set for investors). This part of the thesis is based on Dai, Li and Liu (2009), with the research objectives of:

- To establish a mathematical model in terms of variational inequalities for our finite horizon portfolio selection problem with market closure and transaction costs;
- To examine the effect of transaction costs on liquidity premium;
- To investigate how market closure and return dynamics affect an investor's utility and why daily trading volume is U-shaped in almost all stock exchanges.

The rest of this thesis is organized as follows. Chapter 2 reviews the classical Merton's model for portfolio selection in the absence of transaction costs. This also paves the benchmark for computation of liquidity premium. Chapter 3 is devoted to the finite horizon portfolio selection problem with transaction costs and consumption. We formulate the model, prove the regularity of solution, and characterize the behaviors of trading boundaries. Chapter 4 investigates portfolio selection problem with market closure in the absence of consumption. Liquidity premium is our main concern. Furthermore, we simulate the trading with market opening and closing, and get a U-shaped trading volume pattern, which is consistent with empirical evidence. The last chapter concludes and proposes prospective future research topics.

# Chapter 2

# Merton's finite horizon optimal portfolio selection problem

Merton(1971) pioneered in applying continuous-time stochastic models to study financial markets. He first solved the portfolio selection problem in the absence of transaction costs. His work prepared the foundation for most later research, as well as ours. So we devote this chapter to reviewing Merton's model.

## 2.1 The asset market

Assume the asset market consists of only two investment instruments: one is the risk free asset, which can be a bank account or a government bond; the other asset is risky, which can be a stock. The price processes of the risk-free ($P_t$) and risky ($S_t$) assets are governed by the following SDEs:

$$\begin{aligned} dP_t &= rP_t dt, \\ dS_t &= \mu S_t dt + \sigma S_t dB_t, \end{aligned}$$

where $r > 0$ is the constant risk free interest rate, $\mu > r$ and $\sigma > 0$ are constants, representing the expected return rate and the volatility of the stock return. The process

$\{B_t : t \geq 0\}$ is a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ with $B_0 = 0$ almost surely. We assume $\mathcal{F} = \mathcal{F}_\infty$, the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous and each $\mathcal{F}_t$ contains all null sets of $\mathcal{F}_\infty$.

Assume that a constant relative risk aversion (CRRA) investor holds $X_t$ and $Y_t$ in bank and stock respectively, expressed in monetary terms. In the absence of transaction costs, the equations describing their evolution are

$$dX_t = (rX_t - kC_t)\,dt - dI_t + dD_t \tag{2.1}$$

$$dY_t = \mu Y_t dt + \sigma Y_t dB_t + dI_t - dD_t, \tag{2.2}$$

where $C_t$ is the consumption rate, $I_t$ and $D_t$ are right-continuous (with left hand limits), nonnegative, and nondecreasing $\{\mathcal{F}_t\}_{t\geq 0}$-adapted processes with $I_0 = D_0 = 0$, representing cumulative dollar values for the purpose of buying and selling stock, respectively. Parameter $k$ is taken to be 0 or 1, indicating whether consumption is involved ($k = 1$) or not ($k = 0$). We further assume that consumption withdrawals are made from the bank account.

## 2.2 The investor's problem

When there are no transaction costs, the investor's liquidation wealth at time $t$ can be defined as $W_t = X_t + Y_t$. It is reasonable to require that the net wealth at any time $t$ must always be nonnegative, thus the solvency region $\mathcal{S}$ should be

$$\mathcal{S} = \left\{(x, y) \in R^2 : x + y > 0\right\}.$$

Assume that the investor is given an initial position $(x, y) \in \mathcal{S}$ at time 0. An investment and consumption strategy $(I, D, C)$ is admissible for $(x, y)$ starting from time $s \in [0, T)$ if $(X_t, Y_t)$ given by (2.1)-(2.2 ) with $X_s = x$ and $Y_s = y$ is in $\mathcal{S}$ for all $t \in [s, T]$. We let $\mathcal{A}_s(x, y)$ denote the set of admissible investment strategies starting from time $s$.

The investor's problem is to choose an admissible strategy so as to maximize his expected utility from accumulative consumption and terminal wealth, i.e., to look for

$$\sup_{(I,D,C)\in\mathcal{A}_0(x,y)} E_0^{x,y}\left[\int_0^T e^{-\beta s}U\left(kC_s\right)ds + e^{-\beta T}U(W_T)\right] \tag{2.3}$$

subject to (2.1)-(2.2). Here $\beta > 0$ is discounting rate, $E_t^{x,y}$ denotes the conditional expectation at time $t$ given that time-$t$ endowment $X_t = x$, $Y_t = y$, and the constant relative risk aversion utility function is

$$U(W) = \frac{W^{1-\gamma}-1}{1-\gamma} - \frac{1}{1-\gamma}, \qquad \gamma > 0. \tag{2.4}$$

Note that although (2.4) represents the power utility function, it converges to the logarithm utility function $U(W) = \log(W)$ as $\gamma$ approaches to 1.[1]

## 2.3 The solution in the absence of transaction costs

When no transaction costs apply, we can simplify the investor's problem by redefining the state variable as

$$w \equiv x + y.$$

It is obvious that the wealth process $W_t$ is governed by

$$\begin{aligned} dW_t &= dX_t + dY_t && (2.5)\\ &= [rW_t + (\mu - r)Y_t - C_t]dt + \sigma Y_t dB_t, && (2.6) \end{aligned}$$

then (2.3) can be rewritten as

$$\sup_{(Y,C)\in\mathcal{A}_0(w)} E_0^{w}\left[\int_0^T e^{-\beta s}U\left(kC_s\right)ds + e^{-\beta T}U(W_T)\right]. \tag{2.7}$$

Here, the original stochastic control triple $(I_t, D_t, C_t)$ reduces to $(Y_t, C_t)$.

Define the value function as a function of wealth $w$ and time $t$ by

$$V(w,t) = \sup_{(Y,C)} E_t^{w}\left[\int_t^T e^{-\beta s}U\left(kC_s\right)ds + e^{-\beta T}U(W_T)\right], \tag{2.8}$$

[1] Some literature would like to use "$(1-\gamma)$" as the "$\gamma$" used in this thesis, including Dai and Yi (2009), Dai, Jiang and Yi (2007), and Chong (). In such context, the utility function is defined as $U(W) = \frac{W^\gamma}{\gamma}$, if $\gamma < 1$, $\gamma \neq 0$, and $U(W) = \log(W)$, if $\gamma = 0$.

it satisfies the HJB equation:

$$\max_{Y,C}\left\{V_t+(rw+(\mu-r)\,Y-C)\,V_w+\frac{1}{2}\sigma^2Y^2V_{ww}-\beta V+kU\,(C)\right\}=0 \tag{2.9}$$

with terminal condition

$$V(w,T)=U\,(w)\,.$$

Taking derivative w.r.t. $Y$ and $C$ inside the big parentheses of (2.9), it can be shown that the maximum should be attained at the pair of $(Y_t^*,\ C_t^*)$ which are

$$Y_t^* = -\frac{(\mu-r)\,V_w}{\sigma^2V_{ww}}, \tag{2.10}$$

$$C_t^* = k\,(V_w)^{-\frac{1}{\gamma}}\,. \tag{2.11}$$

Putting them back into the equation can lead to the closed form solution of

$$V(w,t)=\frac{w^{1-\gamma}}{1-\gamma}b^{-\gamma}-\frac{1}{1-\gamma} \tag{2.12}$$

where

$$\begin{aligned} b &= \frac{c}{k-(k-c)e^{-c(T-t)}} \\ c &= \frac{1}{\gamma}\left[\beta-(1-\gamma)\,r-\frac{(1-\gamma)\,(\mu-r)^2}{2\gamma\sigma^2}\right] \end{aligned}$$

To get the optimal strategy, one can put the value function (2.12) back into equation (2.10) and equation (2.11). The optimal trading and consumption (if applicable, i.e. when $k=1$) strategies are

$$Y_t^* = \frac{\mu-r}{\gamma\sigma^2}W_t, \tag{2.13}$$

$$C_t^* = kbW_t. \tag{2.14}$$

Equation (2.13) indicates that in the absence of transaction costs, the optimal trading strategy is to keep a constant fraction of the total wealth in stock, i.e., to maintain

$$\frac{Y_t^*}{W_t}=\frac{\mu-r}{\gamma\sigma^2} \tag{2.15}$$

or in terms of the bank account-to-stock ratio, Merton's optimal strategy is to keep

$$\frac{X_t^*}{Y_t^*} = -\frac{\mu - r - \gamma\sigma^2}{\mu - r}. \tag{2.16}$$

In the rest of this thesis, we will refer to the ratio defined by (2.15) or (2.16) as the so called "Merton line".

Chapter 3

# Finite horizon optimal investment and consumption with transaction costs

Merton's pioneering work in portfolio selection supposes zero transaction costs. The resulting optimal trading strategy is to keep the ratio of stock value to total wealth at a constant level, the Merton line. However, in reality, transaction costs do apply and Merton's strategy would lead to enormous transaction cost payments due to incessant trading. So after Merton's work, the impact of transaction costs has been drawing much attention from researchers. In this chapter, we will investigate the finite horizon portfolio selection problem with transaction costs and consumption.

## 3.1 Problem formulation

We suppose that there are only two assets available for investment: a risk-less asset (bank account) and a risky asset (stock). As in previous chapter, their prices, denoted by $P_t$ and $S_t$, respectively, evolve according to the following equations:

$$\begin{aligned} dP_t &= rP_t dt, \\ dS_t &= \mu S_t dt + \sigma S_t dB_t, \end{aligned}$$

where $r > 0$ is the constant risk-less rate, $\mu > r$ and $\sigma > 0$ are constants, standing for the expected rate of return and the return volatility for the stock, respectively, of the stock. The process $\{B_t;\, t > 0\}$ is a standard Brownian motion on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P\right)$ with $B_0 = 0$ almost surely. We assume $\mathcal{F} = \mathcal{F}_\infty$, the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous and each $\mathcal{F}_t$ contains all null sets of $\mathcal{F}_\infty$.

By now, things are pretty much the same as in previous chapter. However, we are going to see the difference right now.

Assume that a CRRA investor holds $X_t$ and $Y_t$ in bank and stock respectively, expressed in monetary terms. In the presence of transaction costs, the original equations (2.1) and (2.2) governing the evolutions of $X_t$ and $Y_t$ would turn into:

$$dX_t = (rX_t - kC_t)\,dt - (1+\theta)dI_t + (1-\alpha)dD_t \tag{3.1}$$

$$dY_t = \mu Y_t dt + \sigma Y_t dB_t + dI_t - dD_t, \tag{3.2}$$

where $C_t$ is the consumption rate, $I_t$ and $D_t$ are the cumulative stock purchase and sale processes, which are right-continuous (with left hand limits), nonnegative, and non-decreasing $\{\mathcal{F}_t\}_{t\geq 0}$-adapted processes with $I_0 = D_0 = 0$. Parameter $k$ is taken to be 0 or 1, indicating whether consumption is involved ($k = 1$) or not ($k = 0$). The constants $\theta \in (0, \infty)$ and $\alpha \in (0, 1)$ appearing in these equations account for proportional transaction costs incurred on purchase and sale of stock, respectively.

Due to the presence of transaction costs, the investor's net liquidation wealth in monetary terms at time $t$ is

$$W_t = \begin{cases} X_t + (1-\alpha)Y_t & \text{if } Y_t \geq 0, \\ X_t + (1+\theta)Y_t & \text{if } Y_t < 0. \end{cases}$$

Since it is required that the investor's net wealth must be positive, the solvency is decided by

$$\mathcal{S} = \left\{(x, y) \in R^2 : x + (1+\theta)y > 0,\ x + (1-\alpha)y > 0\right\}.$$

Assume that the investor is given an initial position $(x, y) \in \mathcal{S}$ at time 0. An investment and consumption strategy $(I, D, C)$ is admissible for $(x, y)$ starting from time $s \in [0, T)$

if $(X_t, Y_t)$ governed by (3.1)-(3.2) with $X_s = x$ and $Y_s = y$ is in $\mathcal{S}$ for all $t \in [s, T]$. We let $A_s(x, y)$ be the set of all admissible investment strategies starting from time $s$.

The investor's problem is to choose an admissible strategy so as to maximize his expected utility of accumulative consumption and terminal wealth, i.e. to search for

$$\sup_{(I,D,C)\in\mathcal{A}_0(x,y)} E_0^{x,y}\left[\int_0^T e^{-\beta s}U\left(kC_s\right)ds + e^{-\beta T}U(W_T)\right]$$

subject to (3.1)-(3.2). Here again $\beta > 0$ is discounting factor, $E_t^{x,y}$ denotes the conditional expectation at time $t$ given that initial endowment $X_t = x$, $Y_t = y$, and the utility function is

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma} - \frac{1}{1-\gamma},$$

and $\gamma > 0$ is the constant relative risk aversion coefficient.

Note that when $k = 0$, no consumption is involved and the investor only aims at maximizing the expected utility of terminal wealth; while for the case $k = 1$, the investor derives utility from intermediate consumption in addition to terminal wealth.

We define the value function by

$$V(x,y,t) = \sup_{(I,D,C)\in\mathcal{A}_t(x,y)} E_t^{x,y}\left[\int_t^T e^{-\beta(s-t)}U\left(kC_s\right)ds + e^{-\beta(T-t)}U(W_T)\right], \quad (x,y)\in\mathcal{S},\ t\in[0,T).$$

It then satisfies the following Hamilton-Jacobi-Bellman equation:

$$\min\left\{-V_t - \mathscr{L}V, -(1-\alpha)V_x + V_y, (1+\theta)V_x - V_y\right\} = 0, \quad (x,y)\in\mathcal{S},\ t\in[0,T) \tag{3.3}$$

with the terminal condition

$$V(x,y,T) = \begin{cases} U\left(x+(1-\alpha)y\right) & \text{if } y \geq 0, \\ U\left(x+(1+\theta)y\right) & \text{if } y < 0, \end{cases} \tag{3.4}$$

where

$$\mathscr{L}V = \frac{1}{2}\sigma^2 y^2 V_{yy} + \mu y V_y + rxV_x - k\beta V + k\frac{\gamma}{1-\gamma}\left(V_x\right)^{-\frac{1-\gamma}{\gamma}}$$

The problem of portfolio selection in the absence of consumption ($k = 0$) has been thoroughly studied by Dai and Yi (2009). In the rest of this chapter, we will focus on the consumption case and carry on studying system (3.3) with $k = 1$.

### 3.1.1 A variational inequality with gradient constraints

The homogeneity of the utility function and the fact that $\mathcal{A}(\rho x, \rho y) = \rho\mathcal{A}(x, y)$ for all $\rho > 0$ imply that $V + \frac{1}{1-\gamma}$ is concave in $(x, y)$ and homogeneous of degree $1-\gamma$ in $(x, y)$ [cf. Fleming and Soner (1993), Lemma VIII.3.2]. Say,

$$V(\rho x, \rho y, t) + \frac{1}{1-\gamma} = \rho^{1-\gamma}\left(V(x, y, t) + \frac{1}{1-\gamma}\right), \ \forall \rho > 0.$$

This inspires us to take the following transformation[1]:

$$\begin{aligned} \phi(\frac{x}{y}, t) &= V(\frac{x}{y}, 1, t) + \frac{1}{1-\gamma} && (3.5) \\ &= \left(\frac{1}{y}\right)^{1-\gamma} \cdot \left[V(x, y, t) + \frac{1}{1-\gamma}\right]. && (3.6) \end{aligned}$$

After transformation, our value function becomes $\phi(\cdot, \cdot) : (\alpha - 1, \infty) \times [0, T) \longrightarrow \mathbb{R}$. For the purpose of saving notations, we still denote the state variable in $\phi(\cdot, \cdot)$ by $x$, then the governing equation for $\phi(x, t)$ is given by

$$\begin{cases} \min\Big\{-\phi_t - \mathcal{L}_1\phi - \frac{\gamma}{1-\gamma}(\phi_x)^{-\frac{1-\gamma}{\gamma}}, \\ \qquad\qquad -(x+1-\alpha)\phi_x + (1-\gamma)\phi, \ (x+1+\theta)\phi_x - (1-\gamma)\phi\} = 0, \\ \phi(x, T) = \frac{1}{1-\gamma}(x+1-\alpha)^{1-\gamma}, \quad -(1-\alpha) < x < +\infty, \ 0 \le t < T, \end{cases} \tag{3.7}$$

where

$$\mathcal{L}_1\phi \equiv \frac{1}{2}\sigma^2 x^2 \phi_{xx} + \beta_2 x\phi_x + \beta_1\phi - \beta\phi$$

with $\beta_2 = -\left(\mu - r - \gamma\sigma^2\right)$, $\beta_1 = \mu - \frac{1}{2}\gamma\sigma^2$, and all parameters $\beta$, $\mu$, $r$, $\sigma$, $\gamma$, $\alpha$ and $\theta$ are constant, $\beta > 0$, $\mu > r > 0$, $\sigma > 0$, $\theta \in [0, \infty)$, $\alpha \in [0, 1)$, $\alpha + \theta > 0$, $\gamma > 0$.

Similar to Dai and Yi (2009), we further make use of the transformation

$$w(x, \tau) = \frac{1}{1-\gamma}\ln\left[(1-\gamma)\,\phi(x, t)\right]$$

where

$$\tau = T - t,$$

[1] Here we only consider the transformation in the region where $y > 0$. Actually, we can show $\{y \le 0\}$ is always in the buying region (to be defined later in this thesis) by using similar arguments as in Shreve and Soner (1994) or Dai and Yi (2009), so this simplification will not cause loss of generality.

then problem (3.7) can be reduced to a parabolic variational inequality with gradient constraints.

**Problem A:**

$$\begin{cases} \min \left\{ w_\tau - \mathcal{L}_2 w - \frac{\gamma}{1-\gamma} \left(e^w w_x\right)^{-\frac{1-\gamma}{\gamma}}, \ \frac{1}{x+1-\alpha} - w_x, \ w_x - \frac{1}{x+1+\theta} \right\} = 0, \\ w(x,0) = \ln\left(x+1-\alpha\right), \qquad -\left(1-\alpha\right) < x < +\infty, \, 0 < \tau \leq T. \end{cases}$$

here

$$\mathcal{L}_2 w = \frac{1}{2}\sigma^2 x^2 \left( w_{xx} + (1-\gamma)\left(w_x\right)^2 \right) + \beta_2 x w_x + \beta_1 - \frac{1}{1-\gamma}\beta.$$

PDE problems related to problem A (variational inequalities with gradient constraints) have been studied by many researchers, including Evans (1979), Wiegner (1981), Ishii and Koike (1983), Hu (1986), Soner and Shreve (1991) and Zhu (1992). It is Evans (1979) who first considered this type (elliptic) problem and showed that the solution to this type problem has a solution in $W^{1,p} \cap W^{2,p}_{loc}$ $(1 \leq p < \infty)$. This regularity turns out to be sharp in the absence of convexity. But, the present problem does have the convexity. Hence, we expect better regularity results be available. Indeed, Shreve and Soner (1994) and Dai and Yi (2009) obtained $C^2$ smoothness in the spatial direction for the stationary case and the no-consumption case, respectively. We will show it is still true for the present problem. However, the viscosity solution approach adopted by Shreve and Soner (1994) seems unable to deal with the present time-dependent problem. On the other hand, it is intractable to study the properties of free boundaries directly from problem A. This motivates us to adopt an indirect approach, following Dai and Yi (2009).

### 3.1.2 A double obstacle problem

In the attempt to reduce Problem A to a standard variational inequality, we employ the similar transformation as in Dai and Yi (2009) to define:

$$v \equiv w_x = \frac{1}{1-\gamma}\frac{V_x}{V}. \tag{3.8}$$

Formally we have

$$\begin{aligned}\frac{\partial}{\partial x}\mathcal{L}_2 w &= \frac{1}{2}\sigma^2 x^2 v_{xx} - \left(\mu - r - (1+\gamma)\,\sigma^2\right) x v_x - \left(\mu - r - \gamma\sigma^2\right) v \\ &\qquad + (1-\gamma)\,\sigma^2\left(x^2 v v_x + x v^2\right) \\ &\equiv \mathcal{L}v, \end{aligned} \tag{3.9}$$

and

$$\frac{\partial}{\partial x}\left(-\frac{\gamma}{1-\gamma}\left(e^w w_x\right)^{-\frac{1-\gamma}{\gamma}}\right) = \left(e^{(1-\gamma)w} v\right)^{-\frac{1}{\gamma}}\left(v^2 + v_x\right) \equiv \mathcal{L}_w v. \tag{3.10}$$

Then we postulate that $v$ is the solution to the following standard variational inequality, also termed as double obstacle problem:

$$\begin{cases} \min\left\{\max\left\{v_\tau - \mathcal{L}v + \mathcal{L}_w v, v - \frac{1}{x+1-\alpha}\right\}, v - \frac{1}{x+1+\theta}\right\} = 0, \\ v(x,0) = \frac{1}{x+1-\alpha}, \quad -(1-\alpha) < x < +\infty,\ 0 < \tau \le T, \end{cases}$$

or equivalently,

$$\begin{cases} v_\tau - \mathcal{L}v + \mathcal{L}_w v = 0 & \text{if } \frac{1}{x+1+\theta} < v < \frac{1}{x+1-\alpha}, \\ v_\tau - \mathcal{L}v + \mathcal{L}_w v \le 0 & \text{if } v = \frac{1}{x+1-\alpha}, \\ v_\tau - \mathcal{L}v + \mathcal{L}_w v \ge 0 & \text{if } v = \frac{1}{x+1+\theta}, \\ v(x,0) = \frac{1}{x+1-\alpha}, \quad -(1-\alpha) < x < +\infty,\ 0 < \tau \le T. \end{cases} \tag{3.11}$$

Here $\frac{1}{x+1+\theta}$ and $\frac{1}{x+1-\alpha}$ stand for lower and upper obstacles, respectively. We stress that $v_\tau - \mathcal{L}v + \mathcal{L}_w v \ge 0$ on the lower obstacle and $v_\tau - \mathcal{L}v + \mathcal{L}_w v \le 0$ on the upper obstacle, which has a clear physical interpretation.

It is well known that the solution to a double obstacle problem is of $C^1$ in the spatial direction. We immediately obtain $w \in C^2$ in the spatial direction (but on the degenerate line of $x = 0$) provided that $v = w_x$ satisfies (3.11). Also, we will be able to utilize problem (3.11) to analyze the behaviors of free boundaries.

As a consequence, the main task is to prove that $v = w_x$ is the solution to problem (3.11). Such an idea also appeared in Dai and Yi (2009) to deal with the no consumption case. However, it is not an easy task for the present consumption case because $\mathcal{L}_w v$ depends on $w$ and then problem (3.11) itself is not a self-contained system.

## 3.2 On the double obstacle problem (3.11)

Our final purpose is to prove that the function $v$ defined in (3.8) is the solution to problem (3.11). To achieve this object, we split the arguments into several steps.

### 3.2.1 The problem (3.11) with a known $w(x,\tau)$

Since problem (3.11) is not a self-contained system with the presence of function $w(x,\tau)$ in operator, we would like to first investigate it with a known function $w(x,\tau)$, where $w(x,\tau)$ is assumed to have the following properties:

$$|w(x,\tau)-\ln(x+1-\alpha)| \leq M_T, \tag{3.12}$$

$$\frac{1}{x+1+\theta} \leq w_x(x,\tau) \leq \frac{1}{x+1-\alpha}, \tag{3.13}$$

$$|w_\tau(x,\tau)| \leq M, \tag{3.14}$$

$$w(x,0^+) = \ln(x+1-\alpha). \tag{3.15}$$

Here $M$ and $M_T$ are some positive constants.

Notice that the initial value in (3.11) is unbounded near $x=-(1-\alpha)$. To avoid the trouble of dealing with the unboundedness, we confine problem (3.11) to the domain $\Omega_T=(x^*,+\infty)\times(0,T)$ with a boundary condition

$$v_x\left(x^*,\tau\right)=-\frac{1}{\left(x^*+1-\alpha\right)^2}, \quad \tau\in(0,T)\,, \tag{3.16}$$

where $x^*\in(-(1-\alpha),0)$. We always assume $x^*$ to be close enough to $-(1-\alpha)$. Later we will see that it is without loss of generality.

Still, the domain $\Omega_T$ is unbounded, so we further confine problem (3.11) to a bounded domain $\Omega_T^R=(x^*,R)\times(0,T)$ with $R>0$. On $x=R$ we impose the boundary condition

$$v_x\left(R,\tau\right)+v^2\left(R,\tau\right)=0, \quad \tau\in(0,T)\,. \tag{3.17}$$

Now instead of studying problem (3.11), we study the following problem at the first stage:

$$\begin{cases} v_\tau - \mathcal{L}v + \mathcal{L}_w v = 0, & \text{if } \frac{1}{x+1+\theta} < v < \frac{1}{x+1-\alpha}, \\ v_\tau - \mathcal{L}v + \mathcal{L}_w v \geq 0, & \text{if } v = \frac{1}{x+1+\theta}, \\ v_\tau - \mathcal{L}v + \mathcal{L}_w v \leq 0, & \text{if } v = \frac{1}{x+1-\alpha}, \\ v_x(x^*, \tau) = -\frac{1}{(x^*+1-\alpha)^2}, & \\ v_x(R,\tau) + v^2(R,\tau) = 0, & \qquad\qquad (x,\tau) \in \Omega_T^R \\ v(x,0) = \frac{1}{x+1-\alpha}, & \end{cases} \tag{3.18}$$

Furthermore, since the operator $\mathcal{L}$ is degenerate on $x = 0$, we first consider the following regularized problem with some positive number $\delta > 0$,

$$\begin{cases} (v_\delta)_\tau - \mathcal{L}_\delta v_\delta + \mathcal{L}_w v_\delta = 0, & \text{if } \frac{1}{x+1+\theta} < v_\delta < \frac{1}{x+1-\alpha}, \\ (v_\delta)_\tau - \mathcal{L}_\delta v_\delta + \mathcal{L}_w v_\delta \geq 0, & \text{if } v_\delta = \frac{1}{x+1+\theta}, \\ (v_\delta)_\tau - \mathcal{L}_\delta v_\delta + \mathcal{L}_w v_\delta \leq 0, & \text{if } v_\delta = \frac{1}{x+1-\alpha}, \\ (v_\delta)_x (x^*, \tau) = -\frac{1}{(x^*+1-\alpha)^2}, & \\ (v_\delta)_x (R,\tau) + v_\delta^2(R,\tau) = 0, & \qquad\qquad (x,\tau) \in \Omega_T^R, \\ v_\delta(x,0) = \frac{1}{x+1-\alpha}, & \end{cases} \tag{3.19}$$

where

$$\mathcal{L}_\delta v_\delta = \mathcal{L} v_\delta + \delta \left(v_\delta\right)_{xx}.$$

We have the following proposition regarding system (3.19):

**Proposition 1.** *For a given $w(x,\tau)$ satisfying (3.12)-(3.15) in $\Omega_T^R$, problem (3.19) has a solution $v_\delta \in W_p^{2,1}(\Omega_T^R)$, $1 < p < +\infty$, and*

$$\frac{1}{x+1+\theta} \leq v_\delta \leq \frac{1}{x+1-\alpha} \tag{3.20}$$

$$-\frac{K}{(x+1-\alpha)^2} \leq (v_\delta)_x \leq -v_\delta^2 \tag{3.21}$$

*where $K$ is a positive constant independent of $\delta$ and $R$.*

*Proof.* We use the standard penalty method and the fixed point theorem as in Friedman

(1982), section 1.8 to prove the existence of solution to problem (3.19) in $W_p^{2,1}(\Omega_T^R)$, $1 < p < +\infty$.

Define two penalty functions $\beta_\varepsilon(t)$ and $\gamma_\varepsilon(t)$ as follows:

$$\begin{array}{ll} \beta_\varepsilon(\xi) \le 0, & \gamma_\varepsilon(\xi) \ge 0, \\ \beta_\varepsilon(\xi) = 0 \ \text{ if } \xi \ge \varepsilon, & \gamma_\varepsilon(\xi) = 0 \ \text{ if } \xi \le -\varepsilon, \\ \beta_\varepsilon(0) = -c_1, \ (c_1 > 0), & \gamma_\varepsilon(0) = c_2, \ (c_2 > 0), \\ \beta'_\varepsilon(\xi) \ge 0, & \gamma'_\varepsilon(\xi) \ge 0, \\ \beta''_\varepsilon(\xi) \le 0, & \gamma''_\varepsilon(\xi) \ge 0, \end{array}$$

with $\varepsilon > 0$, and constants $c_1$ and $c_2$ to be chosen later.

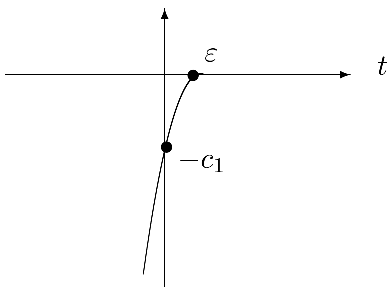



Figure 1: $\beta_\varepsilon(t)$

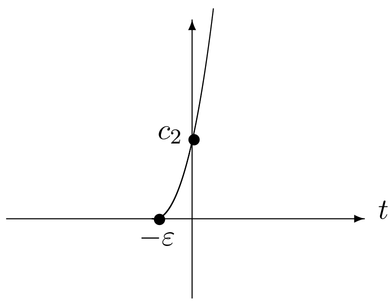



Figure 2: $\gamma_\varepsilon(t)$

For any $\varepsilon > 0$ given, we consider the following approximation problem

$$\left\{ \begin{array}{ll} (v_{\delta,\varepsilon})_\tau - \mathcal{L}_\delta v_{\delta,\varepsilon} + \mathcal{L}_w v_{\delta,\varepsilon} + \beta_\varepsilon(v_{\delta,\varepsilon} - \frac{1}{x+1+\theta}) + \gamma_\varepsilon(v_{\delta,\varepsilon} - \frac{1}{x+1-\alpha}) = 0 & \\ (v_{\delta,\varepsilon})_x(x^*, \tau) = -\frac{1}{(x^*+1-\alpha)^2}, & \\ (v_{\delta,\varepsilon})_x(R, \tau) + v_{\delta,\varepsilon}^2(R, \tau) = 0, & (x, \tau) \in \Omega_T^R \\ v_{\delta,\varepsilon}(x, 0) = \frac{1}{x+1-\alpha}, & \end{array} \right. \tag{3.22}$$

Applying the fixed point theorem, we can prove that problem (3.22) has a solution $v_{\delta,\varepsilon} \in W_p^{2,1}(\Omega_T^R)$. Further, we are able to choose suitable values for $\beta_\varepsilon(0)$ and $\gamma_\varepsilon(0)$ such that $\frac{1}{x+1+\theta}$ and $\frac{1}{x+1-a}$ are subsolution and supersolution of problem (3.22) respectively, namely,

$$\frac{1}{x+1+\theta} \le v_{\delta,\varepsilon} \le \frac{1}{x+1-\alpha} \tag{3.23}$$

This indicates that $\beta_\varepsilon(v_{\delta,\varepsilon} - \frac{1}{x+1+\theta})$ and $\gamma_\varepsilon(v_{\delta,\varepsilon} - \frac{1}{x+1-\alpha})$ are bounded functions, whose bounds are independent of $\varepsilon$. As a consequence,

$$|v_{\delta,\varepsilon}|_{W_p^{2,1}(\Omega_T^R)} \leq C, \tag{3.24}$$

where $C$ is independent of $\varepsilon$. From (3.24) we know that there exists a subsequence of $\{v_{\delta,\varepsilon}\}$ which weakly converges to a $v_\delta$ in $W_p^{2,1}(\Omega_T^R)$, and $v_\delta$ is the solution of problem (3.17).

Letting $\varepsilon \to 0$ in (3.23), (3.20) follows.

Next we prove (3.21).

Clearly

$$(v_\delta)_x + v_\delta^2 = 0, \qquad \text{if } v_\delta = \frac{1}{x+1+\theta} \text{ or } v_\delta = \frac{1}{x+1-\alpha}$$

Then, we only need to show $(v_\delta)_x + v_\delta^2 \leq 0$ in $\mathcal{M}$, where

$$\mathcal{M} = \left\{(x,\tau) \in \Omega_T^R : \frac{1}{x+1+\theta} < v_\delta < \frac{1}{x+1-\alpha}\right\}.$$

Denote $p(x,\tau) = (v_\delta)_x(x,\tau)$ and $q(x,\tau) = v_\delta^2(x,\tau)$, then in $\mathcal{M}$, we have

$$\begin{aligned} & p_\tau - (\frac{1}{2}\sigma^2x^2 + \delta)p_{xx} + (\mu - r - (2+\gamma)\sigma^2)xp_x + (2\mu - 2r - (1+2\gamma)\sigma^2)p \\ & +(e^{(1-\gamma)w}v_\delta)^{-\frac{1}{\gamma}}(q_x + p_x) + \frac{\partial}{\partial x}\left[\left(e^{(1-\gamma)w}v_\delta\right)^{-\frac{1}{\gamma}}\right](p+q) \\ = \; & (1-\gamma)\,\sigma^2(4xv_\delta\,(v_\delta)_x + x^2\left((v_\delta)_x\right)^2 + x^2v_\delta\,(v_\delta)_{xx} + v_\delta^2) \end{aligned} \tag{3.25}$$

and

$$\begin{aligned} & q_\tau - (\frac{1}{2}\sigma^2x^2 + \delta)p_{xx} + (\mu - r - (1+\gamma)\sigma^2)xq_x + 2(\mu - r - \gamma)\sigma^2)p \\ & +2v_\delta\left(e^{(1-\gamma)w}v_\delta\right)^{-\frac{1}{\gamma}}(p+q) \\ = \; & -\sigma^2x^2\left((v_\delta)_x\right)^2 + (1-\gamma)\,\sigma^2\left(2x^2v_\delta^2\,(v_\delta)_x + 2xv_\delta^3\right) - \delta\sigma^2\left((v_\delta)_x\right)^2 \end{aligned}$$

Let

$$H(x,\tau) = p(x,\tau) + q(x,\tau),$$

it is not hard to verify that in $\mathcal{M}$,

$$\begin{aligned}
& H_\tau - (\frac{1}{2}\sigma^2 x^2 + \delta) H_{xx} + (\mu - r - (2+\gamma)\sigma^2 - (1-\gamma)\,\sigma^2 x v_\delta) x H_x \\
& + \left(2\mu - 2r - (1+2\gamma)\,\sigma^2 - 2\,(1-\gamma)\,\sigma^2 x v_\delta\right) H \\
& + (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} H_x + \left\{ \frac{\partial}{\partial x}\left[\left(e^{(1-\gamma)w} v_\delta\right)^{-\frac{1}{\gamma}}\right] + 2 v_\delta (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} \right\} H \\
= \; & -\gamma\sigma^2 (x\,(v_\delta)_x + v_\delta)^2 - \delta\sigma^2 \left((v_\delta)_x\right)^2 \le 0
\end{aligned}$$

And it is straightforward to verify that $H \le 0$ on $\partial\mathcal{M} \cap (\{x = x^*\} \cup \{x = R\} \cup \{\tau = 0\})$. So following the maximum principle (cf. Friedman (1982), p. 74), we then deduce $H \le 0$ in $\mathcal{M}$.

Now we turn to the proof of the left hand side inequality of (3.21). Note that (3.25) can be rewritten as

$$p_\tau - \mathcal{T} p = 0, \quad \text{in } \mathcal{M},$$

where

$$\begin{aligned}
\mathcal{T} p \;=\; & (\frac{1}{2}\sigma^2 x^2 + \delta) p_{xx} - (\mu - r - (2+\gamma)\sigma^2) x p_x - (2\mu - 2r - (1+2\gamma)\sigma^2) p \\
& + (1-\gamma)\,\sigma^2 (x^2 v_\delta p_x + x^2 p^2 + 4 x v_\delta p + v_\delta^2) \\
& + \frac{1}{\gamma}\left(e^{(1-\gamma)w} v_\delta\right)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w} \left((1-\gamma)\, w_x v_\delta + p\right)\left(v_\delta^2 + p\right) \\
& - (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} (2 v_\delta p + p_x)
\end{aligned}$$

It can be verified that for constant $K$ which is big enough,

$$\begin{aligned}
& \left(\frac{\partial}{\partial \tau} - \mathcal{T}\right)\left(-\frac{K}{(x+1-\alpha)^2}\right) \\
= \; & -\frac{1}{(x+1-\alpha)^4}\left[(1-\gamma)\,\sigma^2 x^2 + \frac{1}{\gamma}\left(e^{(1-\gamma)w} v_\delta\right)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w}\right] K^2 \\
& + \frac{1}{(x+1-\alpha)^2}\left\{ \frac{3\sigma^2 x^2 + 6\delta}{(x+1-\alpha)^2} + 2\frac{(\mu - r - (1+2\gamma))\, x - (1-\gamma)\,\sigma^2 x^2 v_\delta}{x+1-\alpha} \right. \\
& \qquad - \left(2\mu - 2r - (1+2\gamma)\,\sigma^2\right) + 4\,(1-\gamma)\,\sigma^2 x v_\delta \\
& \qquad \left. + \left(e^{(1-\gamma)w} v_\delta\right)^{-\frac{1}{\gamma}} \left[\frac{1}{\gamma}\left((1-\gamma)\, w_x + v_\delta\right) + 2\left(\frac{1}{x+1-\alpha} - v_\delta\right)\right]\right\} K \\
& - \frac{1-\gamma}{\gamma}\left(e^{(1-\gamma)w} v_\delta\right)^{-\frac{1}{\gamma}} w_x v_\delta^2 - (1-\gamma)\,\sigma^2 v_\delta^2 \\
\le \; & 0,
\end{aligned}$$

due to the coefficient of the leading term $K^2$ is negative. Here $K$ is independent of $\delta$ and $R$.

It is clear that $p \geq -\frac{K}{(x+1-\alpha)^2}$ on $\partial\mathcal{M} \cap (\{x = x^*\} \cup \{x = R\} \cup \{\tau = 0\})$. Again by the maximum principle, we arrive at the desired result. The proof is complete. $\square$

**Remark 3.2.1.** *The above proof for the left hand side inequality of* (3.21) *requires that* $\gamma < 1$. *Actually, this constraint has been relaxed by Dai and Yang (2009). For the purpose of integrity, we quote the proof of Dai and Yang (2009) in the appendix.*

We are now ready to investigate the properties of function $v(x, \tau)$ in system (3.18). The following proposition holds:

**Proposition 2.** *Under the conditions (3.12)-(3.15), problem (3.16) has a solution* $v \in W_p^{2,1}(\Omega_T^R \backslash \{-\eta < x < \eta\}) \cap C(\overline{\Omega}_T^R)$ *for any small* $\eta > 0$, $1 < p < +\infty$, *and*

$$\frac{1}{x+1+\theta} \leq v \leq \frac{1}{x+1-\alpha} \tag{3.26}$$

$$-\frac{K}{(x+1-\alpha)^2} \leq v_x \leq -v^2 \tag{3.27}$$

$$|v(x,\cdot)|_{C^{\lambda/2}[0,T]} \leq C \tag{3.28}$$

*where* $K$, $\lambda$ *and* $C$ *are positive constants independent of* $R$, *and* $0 < \lambda < 1$.

*Proof.* (3.26)-(3.28) are the consequences of letting $\delta \to 0$ in (3.20)-(3.21) (Let $v$ be the limit of a weakly convergent subsequence of $\{v_\delta\}$ as $\delta \to 0$).

Next, we prove (3.28). When $x > 0$, letting $\delta \to 0$ in (3.17), we infer that (3.16) holds in $\{x > 0\}$ and can be rewritten as

$$\begin{cases} v_\tau - \mathcal{L}v = f(x,\tau), & (x,\tau) \in (0,R) \times (0,T) \\ v_x(R,\tau) + v^2(R,\tau) = 0, & \\ v(x,0) = \frac{1}{x+1-\alpha}, & \end{cases} \tag{3.29}$$

where

$$
\begin{aligned}
f(x,\tau) &= -(e^{(1-\gamma)w}v)^{-\frac{1}{\gamma}}(v^2+v_x) \\
&\quad +\chi_{\{v=\frac{1}{x+1-\alpha}\}}\frac{1}{(x+1-\alpha)^3}[(\mu-r)x+(1-\alpha)(\mu-r-\gamma\sigma^2] \\
&\quad +\chi_{\{v=\frac{1}{x+1+\theta}\}}\frac{1}{(x+1+\theta)^3}[(\mu-r)x+(1+\theta)(\mu-r-\gamma\sigma^2]
\end{aligned}
$$

in which $\chi_A$ is the indicator function on set $A$. Notice that $f(x,\tau)$ is a bounded function, whose bound is independent of $R$.

By transformation

$$x=e^z, \qquad v(x,\tau)=u(z,\tau),$$

then

$$xv_x=u_z, \qquad x^2v_{xx}=u_{zz}-u_z.$$

Thus problem (3.29) becomes

$$
\begin{cases}
u_\tau-\mathcal{L}_z u=g(z,\tau), & (z,\tau)\in(-\infty,\ln R)\times(0,T), \\
u_z(\ln R,\tau)=-Ru^2(\ln R,\tau), & \\
u(z,0)=\frac{1}{e^z+1-\alpha}, &
\end{cases}
\tag{3.30}
$$

where

$$\mathcal{L}_z u=\frac{1}{2}\sigma^2u_{zz}-(\mu-r-(\frac{1}{2}+\gamma)\sigma^2)u_z-(\mu-r-\gamma\sigma^2)u+(1-\gamma)\,\sigma^2\,(e^zu)\,(u_z+u)$$

and $g(z,\tau)=f(e^z,\tau)$ is still a bounded function.

Since $u$ is bounded, the boundary condition $Ru^2\,(\ln R,\tau)$ and the term $e^zu$ appeared in the coefficient of last term of $\mathcal{L}_z u$ are bounded as well, by applying $C^{\lambda,\lambda/2}$ $(0<\lambda<1)$ estimate of parabolic equation, we obtain

$$|u|_{C^{\lambda,\lambda/2}((-\infty,\ln R]\times[0,T])}\le C,$$

where $C$ is independent of $R$. Especially

$$|u(z,\cdot)|_{C^{\lambda/2}[0,T]}\le C.$$

Or, equivalently

$$|v(x,\cdot)|_{C^{\lambda/2}[0,T]} \le C, \qquad 0 \le x \le R.$$

In the same way, we can prove that

$$|v(x,\cdot)|_{C^{\lambda/2}[0,T]} \le C, \qquad x^* \le x \le 0. \tag{3.31}$$

Thanks to (3.27), $v$ is continuous with respect to $x$. This yields, along with (3.31),

$$|v(x,\cdot)|_{C^{\lambda/2}[0,T]} \le C, \qquad x^* \le x \le R.$$

Combining with (3.27) we know that $v \in C(\overline{\Omega}_T^R)$.

At last, we shall prove that $v$ is the solution to (3.16). In fact, we only need to show that (3.16) holds near $x = 0$ in the distributional sense. For any $(0, \tau_0)$ , let us first consider the case

$$\frac{1}{1+\theta} < v\,(0,\tau_0) < \frac{1}{1-\alpha}.$$

Due to the continuity of $v$, there exist $\varepsilon > 0$ and $x_1 < 0 < x_2$, such that

$$\frac{1}{x_2+1+\theta} < v\,(x_2,\tau) < v\,(x_1,\tau) < \frac{1}{x_1+1-\mu}, \quad \text{for } |\tau-\tau_0| < \varepsilon.$$

For fixed $x_1$, $v_\delta\,(x_1,\tau)$ uniformly converges to $v\,(x_1,\tau)$ for $|\tau-\tau_0| < \varepsilon$. So, there is a $\delta_0 > 0$ such that

$$v_\delta\,(x_1,\tau_1) < \frac{1}{x_1+1-\alpha}, \qquad \text{for } |\tau-\tau_0| < \varepsilon, \quad \delta < \delta_0. \tag{3.32}$$

In the same way, for fixed $x_2 > 0$,

$$v_\delta\,(x_2,\tau_2) < \frac{1}{x_2+1+\theta}, \qquad \text{for } \ |\tau-\tau_0| < \varepsilon, \quad \delta < \delta_0. \tag{3.33}$$

Note that (3.32) can be rewritten as

$$(x_1+1-\alpha)^2\, v_\delta\,(x_1,\tau) < x_1+1-\alpha, \quad \text{for } \ |\tau-\tau_0| < \varepsilon, \quad \delta < \delta_0,$$

and

$$\begin{aligned} & \frac{\partial}{\partial x}\Big((x_1+1-\alpha)^2\, v_\delta\,(x_1,\tau) - (x_1+1-\alpha)\Big) \\ = \; & -\left[(x_1+1-\alpha)\, v_\delta\,(x_1,\tau) - 1\right]^2 + (x_1+1-\alpha)^2 \left((v_\delta)_x + v_\delta^2\right) \le 0, \end{aligned} \tag{3.34}$$

where we have used the right hand side inequality in (3.21). We then can deduce

$$(x+1-\alpha)^2 v_\delta(x,\tau) < x+1-\alpha, \quad \text{for} \quad |\tau-\tau_0|<\varepsilon,\ x_1<x<x_2, \quad \delta<\delta_0. \tag{3.35}$$

Namely,

$$v_\delta(x,\tau) < \frac{1}{x+1-\alpha}, \quad \text{for} \quad |\tau-\tau_0|<\varepsilon,\ x_1<x<x_2, \quad \delta<\delta_0 \tag{3.36}$$

On the other hand, we have

$$\frac{\partial}{\partial x}\left(v_\delta - \frac{1}{x+1+\theta}\right) = (v_\delta)_x + \frac{1}{(x+1+\theta)^2} \le (v_\delta)_x + v_\delta^2 \le 0, \tag{3.37}$$

it follows from (3.33) that

$$v_\delta(x,\tau) > \frac{1}{x+1+\theta}, \quad \text{for} \quad |\tau-\tau_0|<\varepsilon,\ x_1<x<x_2, \quad \delta<\delta_0 \tag{3.38}$$

From (3.36) and (3.38), we infer that the first equation of (3.19) holds in

$$E \equiv \{x_1<x<x_2,\ |\tau-\tau_0|<\varepsilon\}.$$

We then deduce by letting $\delta \to 0$ that the first equation of (3.11) holds in $E$ in the distributional sense.

Now we move on to the case of

$$v_\delta(0,\tau_0) = \frac{1}{1+\theta}.$$

Note that

$$v_\delta(0,\tau_0) = \frac{1}{1-\alpha},$$

using a similar argument, we deduce that there is a neighborhood $E$ of $(0,\tau_0)$, such that

$$v_\delta(x,\tau) < \frac{1}{x+1-\alpha}, \quad (x,\tau)\in E$$

when $\delta$ is sufficiently small. Then we have

$$(v_\delta)_\tau - \mathcal{L}_\delta v_\delta + \mathcal{L}_w v_\delta \ge 0 \text{ in } E.$$

Again, we let $\delta \to 0$ to get the desired result. The case of $v(0,\tau_0) = \frac{1}{1-\mu}$ is similar. The proof is complete.

□

Thanks to the right hand side inequality in (3.27), we infer that (3.34) and (3.37) are valid for $v$. This indicates that there are two functions $x_{s,w}(\tau)$ and $x_{b,w}(\tau)$, $\tau \in (0,T]$ such that

$$\left\{(x,\tau) \in \Omega_T^R : v = \frac{1}{x+1-\alpha}\right\} = \left\{(x,t) \in \Omega_T^R,\ x \le x_{s,w}(\tau)\right\}, \tag{3.39}$$

$$\left\{(x,\tau) \in \Omega_T^R : v = \frac{1}{x+1+\theta}\right\} = \left\{(x,t) \in \Omega_T^R,\ x \ge x_{b,w}(\tau)\right\}. \tag{3.40}$$

Regarding the functions $x_{s,w}(\tau)$ and $x_{b,w}(\tau)$, we have the following proposition, which prepares the foundation for our later argument regarding equivalence between Problem A and Problem B (to be defined in next section).

**Proposition 3.** *Denote*

$$x_M = -\frac{\mu - r - \gamma\sigma^2}{\mu - r}$$

*and assume that*

$$x^* \in \left(-(1-\alpha),\ (1-\alpha)\, x_M\right)$$

*Then*

$$x_{s,w}(\tau) \le x_{s,w}\left(0^+\right) \equiv \lim_{\tau \to 0^+} x_{s,w}(\tau) = (1-\alpha)\, x_M. \tag{3.41}$$

*Moreover,* $x_{s,w}(\tau) \in C^\infty$ *when* $x_{s,w}(\tau) > x^*$.

*Proof.* We start from proving (3.41). Note that for $\forall x < x_{s,w}(\tau)$,

$$\begin{aligned} 0 &\ge \left(\frac{\partial}{\partial_\tau} - \mathcal{L} + \mathcal{L}_w\right)\left(\frac{1}{x+1-\alpha}\right) \\ &= \frac{1-\alpha}{(x+1-\alpha)^3}\left[(\mu - r)\, x + (1-\alpha)\left(\mu - r - \gamma\sigma^2\right)\right] \end{aligned}$$

from which we immediately infer that

$$x_{s,w}(\tau) \le -\frac{\mu - r - \gamma\sigma^2}{\mu - r}(1-\alpha) = (1-\alpha)\, x_M.$$

To show that

$$x_{s,w}\left(0^+\right) = (1-\alpha)\, x_M$$

let us first suppose not, then it follows

$$x_{s,w}\left(0^+\right) < (1-\alpha)\, x_M.$$

For any $x_{0^+} \in (x_{s,w}(0), \ (1-\alpha)\, x_M)$, applying the equation

$$v_\tau - \mathcal{L}v + \mathcal{L}_w v = 0$$

at $\tau = 0$ gives

$$v_\tau|_{\tau=0,x=x_0} = \mathcal{L}v - \mathcal{L}_w v|_{\tau=0,x=x_0} = \mathcal{L}\left(\frac{1}{x+1-\alpha}\right)\Big|_{x=x_0} > 0$$

which conflicts with the apparent fact

$$v_\tau|_{\tau=0} \leq 0.$$

So it must be satisfied that

$$x_{s,w}\left(0^+\right) = (1-\alpha)\, x_M.$$

Using (3.27) and analogous arguments as in Dai, Xu and Zhou (2008), we can show $x_{s,w}(\tau) \in C^\infty$ when $x_{s,w}(\tau) > x^*$. The proof is rather complicated that we would like to put it separately outside of this proposition. □

**Proof of $x_{s,w}(\tau) \in C^\infty$ in Proposition 3**

We split the proof into several steps. First we would like to start with a transformation to simplify the differential operators in (3.11). And we confine the proof to the case of $\mu - r - \gamma\sigma^2 < 0$, when there is a positive number $x^* > 0$ such that the selling boundary $x_{s,w}(\tau)$ lies entirely in the region $\tilde{\Omega} \equiv (x^*, \infty) \times (0, T]$.

Consider the following transformation in $\tilde{\Omega}$ :

$$\begin{cases} z = \ln x \\ u(z,\tau) = \frac{1}{v(x,\tau)} \end{cases} \tag{3.42}$$

Then the following system about function $u(z,\tau)$ is generated from system (3.11):

$$\begin{cases} u_\tau - \mathcal{L}_3 u + \tilde{\mathcal{L}}_w u = 0, \quad e^z + 1 - \alpha < u < e^z + 1 + \theta \\ u_\tau - \mathcal{L}_3 u + \tilde{\mathcal{L}}_w u \geq 0, \quad u = e^z + 1 - \alpha \\ u_\tau - \mathcal{L}_3 u + \tilde{\mathcal{L}}_w u \leq 0, \quad u = e^z + 1 + \theta \\ u(z,0) = e^z + 1 - \alpha \end{cases} \tag{3.43}$$

where

$$\mathcal{L}_3 u = \frac{1}{2}\sigma^2 u_{zz} - \left(\mu - r - (\frac{1}{2} + \gamma)\sigma^2\right) u_z + \left(\mu - r - \gamma\sigma^2\right) u - \sigma^2 \frac{u_z^2}{u} + (1-\gamma)\,\sigma^2 e^z \left(\frac{u_z}{u} - 1\right)$$

and

$$\tilde{\mathcal{L}}_w u = \left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}} \left(e^{-z} u_z - 1\right).$$

The original solvency region $\tilde{\Omega}$ is transformed to $\Omega_1 \equiv (\log(x^*),\ \infty) \times (0, T]$, with the corresponding selling, buying and no-transaction region in the $(z, \tau)$ plane characterized by:

$$\begin{aligned}
\mathbf{SR}_z &= \left\{(z,\tau) \in \Omega_1 : u(z,\tau) = e^z + 1 - \alpha\right\},\\
\mathbf{BR}_z &= \left\{(z,\tau) \in \Omega_1 : u(z,\tau) = e^z + 1 + \theta\right\},\\
\mathbf{NT}_z &= \left\{(z,\tau) \in \Omega_1 : e^z + 1 - \alpha < u(z,\tau) < e^z + 1 + \theta\right\}.
\end{aligned}$$

Before proving the smoothness of the selling boundary, we first provide some lemmas about certain estimates of function $u\,(z, \tau)$.

**Lemma 3.2.2.** *Let $u\,(z,\tau)$ be the solution to the double obstacle problem* (3.43)*, then*

*(i) For $\forall z$, it holds that*

$$u_z \geq e^z; \tag{3.44}$$

*(ii) There exists a constant $C_1 > 0$ such that*

$$|u - u_z| \leq C_1. \tag{3.45}$$

*Proof.* Part (i) is the immediate conclusion from

$$v_x + v^2 \leq 0,$$

which is equivalent to

$$\frac{1 - u_z e^{-z}}{u^2} \leq 0, \quad \text{or} \quad u_z \geq e^z.$$

Part (ii) is equivalent to

$$\left|\frac{v + x v_x}{v^2}\right| \leq C_1,$$

which can be derived from $v_x \geq -\frac{K}{(x+1-\alpha)^2}$ and the fact that domain $\Omega_1$ is bounded from below by $\log(x^*)$. □

**Lemma 3.2.3.** *There is a constant $C_2 > 0$, such that $|u_\tau| \leq C_2$.*

*Proof.* Again we consider in the no-transaction region. Note that from

$$\frac{\partial}{\partial \tau}\left(u_\tau - \mathcal{L}_3 u + \tilde{\mathcal{L}}_w u\right) = 0,$$

we have

$$\begin{aligned} &\left[\frac{\partial}{\partial \tau} - \mathcal{L}_4 + e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\frac{\partial}{\partial z} + \frac{e^{-z}\left(u_z - e^z\right)}{\gamma u}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\right](u_\tau) \\ = \;& \frac{1}{\gamma}e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[\left(u_z - e^z\right)\left(\left(1-\gamma\right)w_\tau\right)\right], \end{aligned}$$

where the right-hand-side of the PDE is bounded. Also notice that on the trading boundaries, $u_\tau = 0$ while

$$\begin{aligned} u_\tau|_{\tau=0} &= \left(\mathcal{L}_3 - \tilde{\mathcal{L}}_w\right)\left(e^z + 1 - \mu\right) \\ &= \mathcal{L}_3\left(e^z + 1 - \mu\right) \end{aligned}$$

is bounded, we can construct auxiliary functions in the form of $K_1 e^{K_2 \tau}$, and obtain the boundedness of $u_\tau$ by comparison principle. The proof is complete. □

We are now to prove that $z_{s,w}(\cdot)$ is $C^\infty$, where $z_{s,w}(\tau) = \log\left(x_{s,w}(\tau)\right)$. Thanks to the bootstrap technique, we only need to show that $z_{s,w}(\cdot)$ is Lipschitz-continuous. Then it suffices to prove the cone property, namely, for any $(z, \tau) \in (\log(x^*), \infty) \times (0, T)$, there exists a constant $C > 0$ such that

$$\pm \tau u_\tau - C\frac{\partial}{\partial z}\left[u - \left(e^z + 1 - \mu\right)\right] \leq 0,$$

which is equivalent to

$$C\left(u_z - e^z\right) - \tau u_\tau \geq 0, \tag{3.46}$$

$$C\left(u_z - e^z\right) + \tau u_\tau \geq 0. \tag{3.47}$$

It is easy to check that (3.46) and (3.47) holds in $BR_z$ and $SR_z$. Denote

$$p(z, \tau) = u_z,$$

then we focus the discussion in $NT_z$, in which we have

$$u_\tau - \mathcal{L}_3 u + \tilde{\mathcal{L}}_w u = 0, \tag{3.48}$$

with

$$\mathcal{L}_3 u = \frac{1}{2}\sigma u_{zz} - \left(\mu - r - (\frac{1}{2}+\gamma)\sigma^2\right) u_z + \left(\mu - r - \gamma\sigma^2\right) u - \sigma^2 \frac{u_z^2}{u} + (1-\gamma)\,\sigma^2 e^z \left(\frac{u_z}{u} - 1\right).$$

Taking partial derivative regarding $z$ leads to

$$\frac{\partial}{\partial z}\left(u_\tau - \mathcal{L}_3 u + \tilde{\mathcal{L}}_w u\right) = \left(\frac{\partial}{\partial \tau} - \mathcal{L}_4\right) p - (1-\gamma)\,\sigma^2 e^z \left(\frac{u_z}{u} - 1\right) + \frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right) = 0,$$

or equivalently,

$$\left(\frac{\partial}{\partial \tau} - \mathcal{L}_4\right) p = (1-\gamma)\,\sigma^2 e^z \left(\frac{u_z}{u} - 1\right) - \frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right), \tag{3.49}$$

where

$$\begin{aligned} \mathcal{L}_4 \;=\;& \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial z^2} - \left(\mu - r - (\frac{1}{2}+\gamma)\sigma^2\right)\frac{\partial}{\partial z} + \left(\mu - r - \gamma\sigma^2\right) \\ & -\sigma^2\left(\frac{2u_z - (1-\gamma)\,e^z}{u}\right)\frac{\partial}{\partial z} + \sigma^2\left(\frac{u_z^2 - (1-\gamma)\,e^z u_z}{u^2}\right). \end{aligned}$$

Then by (3.49) it is straightforward to get

$$\begin{aligned} & \left(\frac{\partial}{\partial \tau} - \mathcal{L}_4\right)(u_z - e^z) \\ =\;& (1-\gamma)\,\sigma^2 e^z \left(\frac{u_z}{u} - 1\right) - \frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right) + \mathcal{L}_4 e^z, \end{aligned}$$

while

$$\begin{aligned} \mathcal{L}_4 e^z \;&=\; \sigma^2 e^z \left[\left(1 - \frac{u_z}{u}\right)^2 + \frac{(1-\gamma)\,e^z}{u}\left(1 - \frac{u_z}{u}\right)\right] \\ &=\; \sigma^2 e^z \left(1 - \frac{u_z}{u}\right)\left[1 - \frac{u_z}{u} + \frac{(1-\gamma)\,e^z}{u}\right]. \end{aligned}$$

So

$$\begin{aligned} & \left(\frac{\partial}{\partial \tau} - \mathcal{L}_4\right)(u_z - e^z) \\ =\;& \sigma^2 e^z \left(1 - \frac{u_z}{u}\right)\left[\frac{\gamma u - u_z + (1-\gamma)\,e^z}{u}\right] - \frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right) \\ =\;& \sigma^2 e^z \left(1 - \frac{u_z}{u}\right)\left[\frac{\gamma\,(u - e^z) - (u_z - e^z)}{u}\right] - \frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right) \\ =\;& \gamma\frac{\sigma^2 e^z}{u^2}\,(u - u_z)\,(u - e^z) - \frac{\sigma^2 e^z}{u^2}\,(u - u_z)\,(u_z - e^z) - \frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right). \end{aligned} \tag{3.50}$$

Note that

$$\begin{aligned}\frac{\partial}{\partial z}\left(\tilde{\mathcal{L}}_w u\right) &= e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[u_{zz}-u_z-\frac{1}{\gamma}\left(u_z-e^z\right)\left(\left(1-\gamma\right)w_z-\frac{u_z}{u}\right)\right] \\ &= e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[\frac{\partial}{\partial z}\left(u_z-e^z\right)\right] \\ &\quad -e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[1+\frac{1}{\gamma}\left(\left(1-\gamma\right)w_z-\frac{u_z}{u}\right)\right]\left(u_z-e^z\right),\end{aligned} \tag{3.51}$$

denote

$$\begin{aligned}\mathcal{L}_5 &= \mathcal{L}_4-\frac{\sigma^2 e^z}{u^2}\left(u-u_z\right)-e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\frac{\partial}{\partial z} \\ &\quad +e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[1+\frac{1}{\gamma}\left(\left(1-\gamma\right)w_z-\frac{u_z}{u}\right)\right],\end{aligned} \tag{3.52}$$

then (3.51) is equivalent to

$$\left(\frac{\partial}{\partial \tau}-\mathcal{L}_5\right)\left(u_z-e^z\right)=\gamma\frac{\sigma^2 e^z}{u^2}\left(u-u_z\right)\left(u-e^z\right). \tag{3.53}$$

Also note from

$$\frac{\partial}{\partial \tau}\left(u_\tau-\mathcal{L}_3 u+\tilde{\mathcal{L}}_w u\right)=0,$$

we have

$$\left(\frac{\partial}{\partial \tau}-\mathcal{L}_4\right)\left(u_\tau\right)=-\frac{\partial}{\partial \tau}\left(\tilde{\mathcal{L}}_w u\right), \tag{3.54}$$

which leads to

$$\left(\frac{\partial}{\partial \tau}-\mathcal{L}_4\right)\left(\tau u_\tau\right)=u_\tau+\tau\left(\frac{\partial}{\partial \tau}-\mathcal{L}_4\right)\left(u_\tau\right)=u_\tau-\tau\frac{\partial}{\partial \tau}\left(\tilde{\mathcal{L}}_w u\right). \tag{3.55}$$

Meanwhile,

$$\frac{\partial}{\partial \tau}\left(\tilde{\mathcal{L}}_w u\right)=e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[u_{z\tau}-\frac{1}{\gamma}\left(u_z-e^z\right)\left(\left(1-\gamma\right)w_\tau-\frac{u_\tau}{u}\right)\right], \tag{3.56}$$

so from (3.55) we know that

$$\begin{aligned}
&\left(\frac{\partial}{\partial\tau}-\mathcal{L}_4\right)(\tau u_\tau) \\
= \;& u_\tau - e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[\frac{\partial}{\partial z}(\tau u_\tau)-\frac{1}{\gamma}\left(w_z-\frac{u_z}{u}+\gamma\right)(\tau u_\tau)\right] \\
& -e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\frac{1}{\gamma}\left(w_z-\frac{u_z}{u}+\gamma\right)(\tau u_\tau) \\
& +\tau e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[\frac{1}{\gamma}\left(u_z-e^z\right)\left((1-\gamma)w_\tau-\frac{u_\tau}{u}\right)\right],
\end{aligned} \tag{3.57}$$

i.e.

$$\begin{aligned}
&\left(\frac{\partial}{\partial\tau}-\mathcal{L}_5\right)(\tau u_\tau) \\
= \;& u_\tau+\frac{\sigma^2 e^z}{u^2}(u-u_z)(\tau u_\tau) \\
& -e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\frac{1}{\gamma}\left(w_z-\frac{u_z}{u}+\gamma\right)(\tau u_\tau) \\
& +\tau e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[\frac{1}{\gamma}\left(u_z-e^z\right)\left((1-\gamma)w_\tau-\frac{u_\tau}{u}\right)\right].
\end{aligned} \tag{3.58}$$

The combination of (3.53) and (3.58) gives

$$\begin{aligned}
&\left(\frac{\partial}{\partial\tau}-\mathcal{L}_5\right)\left[C\left(u_z-e^z\right)-\tau u_\tau\right] \\
= \;& C\gamma\frac{\sigma^2 e^z}{u^2}(u-u_z)(u-e^z) \\
& -u_\tau-\frac{\sigma^2 e^z}{u^2}(u-u_z)(\tau u_\tau) \\
& +e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\frac{1}{\gamma}\left(w_z-\frac{u_z}{u}+\gamma\right)(\tau u_\tau) \\
& -\tau e^{-z}\left(\frac{e^{(1-\gamma)w}}{u}\right)^{-\frac{1}{\gamma}}\left[\frac{1}{\gamma}\left(u_z-e^z\right)\left((1-\gamma)w_\tau-\frac{u_\tau}{u}\right)\right].
\end{aligned} \tag{3.59}$$

Taking one more step to define another differential operator

$$\mathcal{J}=\mathcal{L}_5-(1-\gamma)\frac{\sigma^2 e^z}{u},$$

we derive from (3.59) that

$$
\begin{aligned}
&\left(\frac{\partial}{\partial\tau}-\mathcal{J}\right)\left[C\left(u_z-e^z\right)-\tau u_\tau\right] \\
=\;& \left(\frac{\partial}{\partial\tau}-\mathcal{L}_5\right)\left[C\left(u_z-e^z\right)-\tau u_\tau\right]+(1-\gamma)\frac{\sigma^2 e^z}{u}\left[C\left(u_z-e^z\right)-\tau u_\tau\right] \\
=\;& C(1-\gamma)\frac{\sigma^2 e^z}{u^2}\left(u-e^z\right)^2+C(1-\gamma)\frac{\sigma^2 e^{2z}}{u^2}\left(u_z-e^z\right) && (3.60)\\
&-u_\tau-\frac{\sigma^2 e^z}{u^2}\left(u-u_z\right)\left(\tau u_\tau\right)-(1-\gamma)\frac{\sigma^2 e^z}{u}\left(\tau u_\tau\right) && (3.61)\\
&+e^{-z}\left(\frac{e^{\gamma w}}{u}\right)^{-\frac{1}{1-\gamma}}\frac{1}{1-\gamma}\left(w_z-\frac{u_z}{u}+1-\gamma\right)\left(\tau u_\tau\right) && (3.62)\\
&-\tau e^{-z}\left(\frac{e^{\gamma w}}{u}\right)^{-\frac{1}{1-\gamma}}\left[\frac{1}{1-\gamma}\left(u_z-e^z\right)\left(\gamma w_\tau-\frac{u_\tau}{u}\right)\right]. && (3.63)
\end{aligned}
$$

It is not hard to check that all terms in (3.61)-(3.63) are bounded, so we infer that there is a positive constant $C_3$, such that

$$\left|(3.61)+(3.62)+(3.63)\right| \leq C_3.$$

Furthermore, we are very fortunate to have the uniform bound of $u-e^z$, which is

$$u-e^z \geq 1-\alpha>0.$$

So from (3.60)-(3.63) we assert that

$$\left(\frac{\partial}{\partial\tau}-\mathcal{J}\right)\left[C\left(u_z-e^z\right)-\tau u_\tau\right] \geq C(1-\gamma)(1-\alpha)^2\sigma^2\frac{e^z}{u^2}-C_3, \quad \text{in } NT_z.$$

Since the no-transaction region is unbounded, we can follow Shreve and Soner (1991) to introduce an auxiliary function

$$\psi\left(z,\tau;z_0\right)=e^{a\tau}\left(z-z_0\right)^2$$

with a constant $a>0$. We can choose $a$ big enough such that

$$\left(\frac{\partial}{\partial\tau}-\mathcal{J}\right)\psi\left(z,\tau;z_0\right) \geq C_4\left(z-z_0\right)^2-C_5,$$

where $C_4$ and $C_5$ are positive constants independent of $(z,\tau)$. It follows

$$
\begin{aligned}
&\left(\frac{\partial}{\partial\tau}-\mathcal{J}\right)\left[C\left(u_z-e^z\right)-\tau u_\tau+\psi\left(z,\tau;z_0\right)\right] \\
\geq\;& C(1-\gamma)(1-\alpha)^2\sigma^2\frac{e^z}{u^2}-C_3+C_4\left(z-z_0\right)^2-C_5,
\end{aligned}
$$

then we can choose $r > 0$ such that

$$C_4 r^2 - C_3 - C_5 \geq 0,$$

and choose $C > 0$ big enough such that

$$C\left(1-\gamma\right)\left(1-\alpha\right)^2 \sigma^2 \frac{e^z}{u^2} - C_3 - C_5 \geq 0 \quad \text{for } \left|z - z_0\right| \leq r.$$

It then follows

$$\left(\frac{\partial}{\partial \tau} - \mathcal{J}\right)\left[C\left(u_z - e^z\right) - \tau u_\tau + \psi\left(z, \tau; z_0\right)\right] \geq 0, \text{ in } NT_z.$$

Applying the maximum principle and penalty approximation, we conclude

$$C\left(u_z - e^z\right) - \tau u_\tau + \psi\left(z, \tau; z_0\right) \geq 0, \quad (z, \tau) \in (\log(x^*), \infty) \times (0, T).$$

Letting $z = z_0$, we get the desired result (3.46).

The proof of (3.47) can be done in the similar way. So the selling boundary $z_{s,w}(\tau)$ is in $C^\infty$. Back to the $(x, \tau)$ plane, we know that the function $x_{s,w}(\tau) \in C^\infty$.

### 3.2.2 The problem (3.11) with an auxiliary condition

As mentioned before, we need an auxiliary condition to make problem (3.11) self-contained. Now, let us exploit the condition in this section.

Assume $v = w_x$ is a solution to problem (3.11). Due to Proposition 3, we expect that there is a function $x_s\left(\tau\right) : [0, T) \to \left(-\left(1-\alpha\right), +\infty\right)$, such that

$$\left\{(x, \tau) \in \Omega_T : v = \frac{1}{x + 1 - \alpha}\right\} = \left\{(x, \tau) \in \Omega_T : x \leq x_s\left(\tau\right)\right\}.$$

It is apparent that $w(x, \tau) = A(\tau) + \ln\left(x + 1 - \alpha\right)$, $x \leq x_s(\tau)$ with some function $A\left(\tau\right)$ to be determined, while $A\left(0\right) = 0$.

We conjecture that for any $(x, \tau) \in \Omega_T$, it holds that

$$\begin{aligned} w\left(x, \tau\right) &= w(x_s(\tau), \tau) + \int_{x_s(\tau)}^{\tau} v(\xi, \tau) d\xi \\ &= A\left(\tau\right) + \ln\left(x_s\left(\tau\right) + 1 - \alpha\right) + \int_{x_s(\tau)}^{x} v\left(\xi, \tau\right) d\xi. \end{aligned}$$

It is expected that $v(\cdot,\tau)\in C^1$ and $w(\cdot,\tau)\in C^2$. Thus, we should have

$$\begin{aligned} w_x|_{x=x_s(\tau)} &= \frac{1}{x_s(\tau)+1-\alpha}, \\ w_{xx}|_{x=x_s(\tau)} &= -\frac{1}{(x_s(\tau)+1-\alpha)^2}, \end{aligned}$$

which yields

$$\begin{aligned} A'(\tau) &= w_\tau(x_s(\tau),\tau) = \left.\left(\mathcal{L}_2 w + \frac{\gamma}{1-\gamma} e^{-\frac{1-\gamma}{\gamma}w}(w_x)^{-\frac{1-\gamma}{\gamma}}\right)\right|_{x=x_s(\tau)} \\ &= \frac{\gamma}{1-\gamma} e^{-\frac{1-\gamma}{\gamma}A(\tau)} + f(x_s(\tau)) \end{aligned} \tag{3.64}$$

where

$$f(x_s(\tau)) = \frac{1}{(x_s(\tau)+1-\alpha)^2}\left[r x_s^2(\tau) + (\mu+r)(1-\alpha)x_s(\tau) + (\mu - \frac{1}{2}\gamma\sigma^2)(1-\alpha)^2\right] - \frac{1}{1-\gamma}\beta$$

We note that (3.64) can be rewritten as

$$\left(e^{\frac{1-\gamma}{\gamma}A(\tau)}\right)' = \frac{\gamma}{1-\gamma} f(x_s(\tau))\, e^{\frac{1-\gamma}{\gamma}A(\tau)} + 1,$$

combined with $A(0)=0$, we obtain that

$$\begin{aligned} A(\tau) &= \frac{\gamma}{1-\gamma}\log\left[e^{\frac{1-\gamma}{\gamma}\int_0^\tau f(x_s(\varsigma))d\varsigma} + \int_0^\tau e^{\frac{1-\gamma}{\gamma}\int_{\bar\tau}^\tau f(x_s(\varsigma))d\varsigma} d\bar\tau\right] \\ &\equiv \mathcal{H}(x_s(\tau)) \end{aligned}$$

This gives the auxiliary condition with which we want to combine problem (3.11). Or, in another word, we hope to study the following problem.

**Problem B:** Find $w(x,\tau)$, $v(x,\tau)$ and $x_s(\tau):[0,T)\to(x^*,+\infty)$, such that

(i) $\left\{(x,\tau)\in\Omega_T : v(x,\tau) = \frac{1}{x+1-\alpha}\right\} = \{(x,\tau)\in\Omega_T : x\le x_s(\tau)\}$

(ii) $v(x,\tau)$, $(x,\tau)\in\Omega_T$ satisfies (3.11) in which

$$w(x,\tau) = A(\tau) + \ln(x_s(\tau)+1-\alpha) + \int_{x_s(\tau)}^x v(\xi,\tau)\,d\xi, \tag{3.65}$$

where $A(\tau) = \mathcal{H}(x_s(\tau))$.

Now we are able to establish the following proposition:

**Proposition 4.** *Problem B allows a unique solution* $(w(x,\tau), v(x,\tau), x_s(\tau))$ *satisfying (3.12)-(3.15), (3.26)-(3.28) and (3.41), respectively.*

*Proof.* Once the existence of solution is established, the uniqueness would be apparent. So we focus on the proof of existence, which we will employ the Schauder fixed point theorem to prove.

To begin with, we confined to a bounded domain $\overline{\Omega}_T^R$ and consider the Banach space $\mathcal{B} = C(\overline{\Omega}_T^R)$.

Define

$$\mathcal{D} = \left\{ w(x,\tau) \in \mathcal{B} \middle| \begin{array}{l} |w(x,\tau) - \ln(x+1-\alpha)| \le M_T, \\ \frac{1}{x+1+\theta} \le w_x(x,\tau) \le \frac{1}{x+1-\alpha}, \\ |w_\tau(x,\tau)| \le M,\ w\,(x,0) = \ln\,(x+1-\alpha) \end{array} \right\},$$

where $M$ and $M_T$ are positive constants to be determined, $w_x$ and $w_\tau$ are regarding to weak derivatives. It is clear that $\mathcal{D}$ is a compact convex set in $\mathcal{B}$.

For any given $w(x,\tau) \in \mathcal{D}$, let $v(x,\tau)$ be the solution of problem (3.11) confined to $\overline{\Omega}_T^R$ with boundary conditions (3.16)-(3.17), and $x_{s,w}(\tau)$ be the corresponding free boundary. Now we define a mapping $\mathcal{F} : \mathcal{D} \to \mathcal{B}$ as follows,

$$\mathcal{F}w = \bar{w}(x,\tau) \equiv A(\tau) + \ln(x_{s,w}(\tau) + 1 - \alpha) + \int_{x_{s,w}(\tau)}^{x} v(\xi,\tau)d\xi, \tag{3.66}$$

where $A(\tau) = \mathcal{H}\,(x_s(\tau))$.

In the following we shall prove that $\bar{w}(x,\tau) \in \mathcal{D}$.

By definition, it is obvious that

$$\begin{aligned} \bar{w}(x,0) &= \ln(x+1-\alpha), \\ \bar{w}_x(x,\tau) &= v(x,\tau), \end{aligned}$$

and thus

$$\frac{1}{x+1+\theta} \le \bar{w}_x(x,\tau) \le \frac{1}{x+1-\alpha}.$$

Furthermore, by (3.66) and the fact that $\frac{1}{x+1+\theta} \leq w_x(x,\tau) \leq \frac{1}{x+1-\alpha}$, we have

$$\begin{aligned} & A(\tau) + \ln \frac{x+1+\theta}{x+1-\alpha} + \ln \left( \frac{x_{s,w}(\tau)+1-\alpha}{x_{s,w}(\tau)+1+\theta} \right) \\ \leq\ & \bar{w}(x,\tau) - \ln(x+1-\alpha) \\ \leq\ & A(\tau). \end{aligned} \tag{3.67}$$

So by definition of $A(\tau)$ and by (3.41), $A(\tau)$ is bounded. Then we can deduce that there is a positive constant, denoted by $M_T$ independent of $R$, such that

$$|\bar{w}(x,\tau) - \ln(x+1-\alpha)| < M_T.$$

Now the only thing remains to show is

$$|\bar{w}_\tau(x,\tau)| \leq M. \tag{3.68}$$

By (3.66) , we have

$$\begin{aligned} & w_\tau \\ =\ & A'(\tau) + \int_{x_{s,w}(\tau)}^{x} v(\xi,\tau)d\xi \\ =\ & A'(\tau) + \int_{x_{s,w}(\tau)}^{x} \mathcal{L}v(\xi,\tau)d\xi - \int_{x_{s,w}(\tau)}^{x} \mathcal{L}_w v(\xi,\tau)d\xi \\ =\ & A'(\tau) + \int_{x_{s,w}(\tau)}^{x} \frac{\partial}{\partial \xi}\mathcal{L}_2\bar{w}(\xi,\tau)d\xi - \int_{x_{s,w}(\tau)}^{x} \mathcal{L}_w v(\xi,\tau)d\xi \\ =\ & \frac{\gamma}{1-\gamma}e^{-\frac{1-\gamma}{\gamma}A(\tau)} + \frac{1}{2}\sigma^2x^2\left(v_x + (1-\gamma)v^2\right) + \beta_2 xv \\ & +\beta_1 - \frac{1}{1-\gamma}\beta - \int_{x_{s,w}(\tau)}^{x} \mathcal{L}_w v(\xi,\tau)d\xi \end{aligned} \tag{3.69}$$

Combined with (3.26)-(3.27) and the fact that $A(\tau)$ is bounded, we assert that there is a constant $M_1$ independent of $R$ such that

$$\left| \frac{\gamma}{1-\gamma}e^{-\frac{1-\gamma}{\gamma}A(\tau)} + \frac{1}{2}\sigma^2x^2\left(v_x + (1-\gamma)v^2\right) + \beta_2 xv + \beta_1 - \frac{1}{1-\gamma}\beta \right| \leq M_1$$

Note that $w(x,\tau)$ has a bound depending only on $R$, so we assert that there is a constant $M_2$ only depending on $R$ such that in (3.69)

$$0 \leq -\int_{x_{s,w}(\tau)}^{x} \mathcal{L}_w v(\xi,\tau)d\xi \leq M_2$$

we then set

$$M = M_1 + M_2$$

to obtain (3.68). Thus we have proved that $\bar{w}(x,\tau) \in \mathcal{D}$. And note that $w(x,\tau)$ is arbitrarily given in $\mathcal{D}$, so it is straightforward to state that $\mathcal{F}(\mathcal{D}) \subset \mathcal{D}$.

Owing to the uniqueness of solution, $\mathcal{F}$ must be a one-one mapping. Thanks to the compactness of $\mathcal{D}$, we then infer that $\mathcal{F}$ must be continuous. Applying the Schauder fixed point theorem in the Banach space $\mathcal{B}$ shows that problem B confined to $\overline{\Omega}_T^R$ allows a solution $(w_R(x,\tau), v_R(x,\tau), x_s(\tau))$.

The last step is to extend the result to domain $\overline{\Omega}_T$. We only need to show that $(w_R)_\tau$ has a uniform bound, which is independent of $R$. By the definition of $\mathcal{L}_w$ in (3.10), and the fact that $(w_R)_x = v_R$, we have

$$-\int_{x_s(\tau)}^{x} \mathcal{L}_{w_R} v_R(\xi,\tau)d\xi = \int_{x_s(\tau)}^{x} \frac{\partial}{\partial \xi}\left(\frac{\gamma}{1-\gamma}\left(e^{w_R}v_R\right)^{-\frac{1-\gamma}{\gamma}}\right)d\xi.$$

Combining with (3.69), we obtain

$$(w_R)_\tau(x,\tau) = \frac{1}{2}\sigma^2 x^2\left(v_x + (1-\gamma)v^2\right) + \beta_2 x v + \beta_1 - \frac{1}{1-\gamma}\beta + \frac{\gamma}{1-\gamma}\left(e^{w_R}v_R\right)^{-\frac{1-\gamma}{\gamma}}.$$

As a result, it suffices to show that $e^{w_R}v_R$ has a uniform bound. Similar to (3.67), we have

$$A(\tau) + \ln(x+1+\theta) + \ln\left(\frac{x_s(\tau)+1-\alpha}{x_s(\tau)+1+\theta}\right) \leq w_R(x,\tau) \leq A(\tau) + \ln(x+1-\alpha).$$

Owing to

$$\frac{1}{x+1+\theta} \leq v_R(x,\tau) \leq \frac{1}{x+1-\alpha},$$

we then arrive at

$$\frac{x_s(\tau)+1-\alpha}{x_s(\tau)+1+\theta}e^{A(\tau)} \leq e^{w_R}v_R \leq e^{A(\tau)},$$

which is desired. This completes the proof. □

As the counterpart to $x_{b,w}(\tau)$ in (3.40), we can similarly define the boundary $x_b(\tau)$ related to problem B as follows:

$$\{(x,\tau)\in\Omega_T : x\geq x_b(\tau)\} = \left\{(x,\tau)\in\Omega_T : v=\frac{1}{x+1+\theta}\right\}.$$

### 3.2.3 Equivalence between Problem A and Problem B

Recall that when formulating the optimal investment problem with consumption, we first arrived at a variational inequality with gradient constraints:

**Problem A:**

$$\begin{cases} \min\left\{w_\tau - \mathcal{L}_2 w - \frac{\gamma}{1-\gamma}\left(e^w w_x\right)^{-\frac{1-\gamma}{\gamma}},\ \frac{1}{x+1-\alpha} - w_x,\ w_x - \frac{1}{x+1+\theta}\right\} = 0, \\ w(x,0) = \ln(x+1-\alpha), \qquad -(1-\alpha) < x < +\infty,\ 0<\tau\leq T. \end{cases}$$

where

$$\mathcal{L}_2 w = \frac{1}{2}\sigma^2 x^2\left(w_{xx} + (1-\gamma)(w_x)^2\right) + \beta_2 x w_x + \beta_1 - \frac{1}{1-\gamma}\beta.$$

Then by taking partial derivative of $x$ in $w(x,\tau)$, we formally get another problem:

**Problem B:** Find $w(x,\tau),\ v(x,\tau)$ and $x_s(\tau) : [0,T)\to(x^*,+\infty)$, such that

(i) $\left\{(x,\tau)\in\Omega_T : v(x,\tau) = \frac{1}{x+1-\alpha}\right\} = \{(x,\tau)\in\Omega_T : x\leq x_s(\tau)\}$

(ii) $v(x,\tau),\ (x,\tau)\in\Omega_T$ satisfies (3.11) in which

$$w(x,\tau) = A(\tau) + \ln(x_s(\tau)+1-\alpha) + \int_{x_s(\tau)}^{x} v(\xi,\tau)\,d\xi, \tag{3.70}$$

where $A(\tau) = \mathcal{H}(x_s(\tau))$.

We have investigated the regularity of Problem B in previous subsection; and we devote this subsection to proving the equivalence between Problem A and Problem B. To achieve this objective, we take the solution triple $(v(x,\tau), w(x,\tau), x_s(\tau))$ to Problem B. Since Problem A has a unique viscosity solution, we need only show that $w$ in the triple $(v(x,\tau), w(x,\tau), x_s(\tau))$ is the solution to Problem A.

Define the three regions:

$$\begin{aligned}
\mathbf{SR} &= \left\{(x,\tau)\in\Omega_T : v(x,\tau)=\frac{1}{x+1-\alpha}\right\},\\
\mathbf{BR} &= \left\{(x,\tau)\in\Omega_T : v(x,\tau)=\frac{1}{x+1+\theta}\right\},\\
\mathbf{NT} &= \left\{(x,\tau)\in\Omega_T : \frac{1}{x+1+\theta}<v(x,\tau)<\frac{1}{x+1-\alpha}\right\}.
\end{aligned}$$

In finance, the three regions defined above stand for the selling region, buying region and no-transaction region, respectively. Thanks to Proposition 4, we know that the selling region and the no-transaction region on the $(x,\tau)$ plane are separated by a curve $x_s(\tau)$, i.e.

$$\mathbf{SR}=\{(x,\tau)\in\Omega_T : x\le x_s(\tau)\}\,. \tag{3.71}$$

Similar to (3.39)-(3.40), we infer that there is a function (also a curve) $x_b\,(\tau):(0,T)\to[x^*,+\infty)\cup\infty$, which separates the no-transaction region and the buying region, i.e.,

$$\mathbf{BR}=\{(x,\tau)\in\Omega_T : x\ge x_b(\tau)\}. \tag{3.72}$$

Note that $v=w_x$ satisfies (3.11). Owing to (3.9), we have

$$\begin{aligned}
\frac{\partial}{\partial x}\left(w_\tau-\mathcal{L}_2 w-\frac{\gamma}{1-\gamma}e^{-\frac{1-\gamma}{\gamma}w}\,(w_x)^{-\frac{1-\gamma}{\gamma}}\right) &\le 0,\ w_x=\frac{1}{x+1-\alpha}\quad \text{in } x\le x_s(\tau),\\
\frac{\partial}{\partial x}\left(w_\tau-\mathcal{L}_2 w-\frac{\gamma}{1-\gamma}e^{-\frac{1-\gamma}{\gamma}w}\,(w_x)^{-\frac{1-\gamma}{\gamma}}\right) &= 0\quad \text{in } x_s(\tau)<x<x_b(\tau),\\
\frac{\partial}{\partial x}\left(w_\tau-\mathcal{L}_2 w-\frac{\gamma}{1-\gamma}e^{-\frac{1-\gamma}{\gamma}w}\,(w_x)^{-\frac{1-\gamma}{\gamma}}\right) &\ge 0,\quad w_x=\frac{1}{x+1+\theta}\quad \text{in } x\ge x_b(\tau).
\end{aligned}$$

Note that

$$\left.w_\tau-\mathcal{L}_2 w-\frac{\gamma}{1-\gamma}e^{-\frac{1-\gamma}{\gamma}w}\,(w_x)^{-\frac{1-\gamma}{\gamma}}\right|_{x=x_s(\tau)}=0,$$

we then deduce $w$ is the solution to problem A. Then by Proposition 2 and Proposition 4, we achieve the following theorem:

**Theorem 3.2.4.** *Problem A has a solution $w\,(x,\tau)\in W^{2,1}_\infty(\Omega_T^R)$ for any $R>0$ with $w_x\in C(\overline{\Omega}_T)$ and $w_{xx},\ w_\tau\in L^\infty(\Omega_T)\cap C\left(\overline{\Omega}_T\backslash\{x=0\}\right)$. Moreover, $v=w_x$ satisfies problem (3.11), and*

$$-\frac{K}{(x+1-\alpha)^2}\le v_x\le -v^2, \tag{3.73}$$

$$|w_\tau| \le M, \tag{3.74}$$

*where $K$ and $M$ are positive constants.*

## 3.3 Behaviors of free boundaries

### 3.3.1 Without consumption

Dai and Yi (2009) studied the non-consumption case with power utility function. Let $\bar{v} \equiv v(\cdot; k = 0)$ be the gradient function derived from the same approach as in last section, it satisfies the variational inequalities:

$$\begin{cases} \bar{v}_\tau - \mathcal{L}\bar{v} = 0 & \text{if } \frac{1}{x+1+\theta} < \bar{v} < \frac{1}{x+1-\alpha} \\ \bar{v}_\tau - \mathcal{L}\bar{v} \le 0 & \text{if } \bar{v} = \frac{1}{x+1-\alpha}, \\ \bar{v}_\tau - \mathcal{L}\bar{v} \ge 0 & \text{if } \bar{v} = \frac{1}{x+1+\theta}, \\ \bar{v}(x,0) = \frac{1}{x+1-\alpha}, & (x,\tau) \in \Omega_T \end{cases} \tag{3.75}$$

where $\mathcal{L}$ is the same as defined in (3.9). In contrast to problem (3.11), the nonlinear operator $\mathcal{L}_w$ is absent in (3.75).

Problem (3.75) also gives rise to two free boundaries, denoted by $\bar{x}_s(\tau)$ and $\bar{x}_b(\tau)$, such that the trading regions $\overline{BR}$, $\overline{SR}$, and $\overline{NT}$ in the absence of consumption are characterized by

$$\overline{SR} = \left\{(x,\tau) \in \Omega_T : \bar{v}(x,\tau) = \frac{1}{x+1-\alpha}\right\} = \left\{(x,\tau) \in \Omega_T : x \le \bar{x}_s(\tau)\right\}, \tag{3.76}$$

$$\overline{BR} = \left\{(x,\tau) \in \Omega_T : \bar{v}(x,\tau) = \frac{1}{x+1+\theta}\right\} = \left\{(x,\tau) \in \Omega_T : x \ge \bar{x}_b(\tau)\right\} \tag{3.77}$$

and

$$\begin{aligned} \overline{NT} &= \left\{(x,\tau) \in \Omega_T : \frac{1}{x+1+\theta} < \bar{v}(x,\tau) < \frac{1}{x+1-\alpha}\right\} \\ &= \left\{(x,\tau) \in \Omega_T : \bar{x}_s(\tau) < x < \bar{x}_b(\tau)\right\} \end{aligned} \tag{3.78}$$

The behaviors of $\bar{x}_s(\tau)$ and $\bar{x}_b(\tau)$ are fully investigated in Dai and Yi (2009), which we summarize as follows:

**Proposition 5.** *Let* $\bar{x}_s(\tau)$ *and* $\bar{x}_b(\tau)$ *be the two free boundaries as given in* (3.76) *and* (3.77), *which correspond to problem* (3.75). *Define*

$$\tau_0 = \frac{1}{\mu - r} \log \frac{1+\theta}{1-\alpha} \tag{3.79}$$

$$\tau_1 = \frac{1}{\mu - r - \gamma\sigma^2} \log \frac{1+\theta}{1-\alpha} \tag{3.80}$$

*then* $\bar{x}_s(\tau) < \bar{x}_b(\tau)$, *and*

*(i) both* $\bar{x}_s(\tau)$ *and* $\bar{x}_b(\tau)$ *are monotonically decreasing;*

*(ii) for any* $\tau > 0$,

$$-(1-\alpha) < \lim_{\tau \to \infty} \bar{x}_s(\tau) \le \bar{x}_s(\tau) \le \bar{x}_s\left(0^+\right) = (1-\alpha)\, x_M; \tag{3.81}$$

*Moreover,*

$$\bar{x}_s(\tau) \equiv 0 \quad \textit{if } \mu - r - \gamma\sigma^2 = 0 \tag{3.82}$$

$$\bar{x}_s(\tau) < 0 \quad \textit{if } \mu - r - \gamma\sigma^2 > 0 \tag{3.83}$$

$$\bar{x}_s(\tau) > 0 \quad \textit{if } \mu - r - \gamma\sigma^2 < 0 \tag{3.84}$$

*(iii) for any* $\tau > 0$,

$$\bar{x}_b(\tau) \ge (1+\theta)\, x_M, \tag{3.85}$$

*and*

$$\bar{x}_b(\tau) = +\infty \ \textit{ if and only if } \quad \tau \in (0, \tau_0]; \tag{3.86}$$

*moreover, when* $\mu - r - \gamma\sigma^2 > 0$,

$$\bar{x}_b(\tau) > 0, \ \textit{for } \tau < \tau_1 \tag{3.87}$$

$$\bar{x}_b(\tau_1) = 0 \tag{3.88}$$

$$\bar{x}_b(\tau) < 0, \ \textit{for } \tau > \tau_1. \tag{3.89}$$

**Remark 3.3.1.** *Liu and Loewenstein (2002) obtained partial results of the above proposition, including* (3.81), (3.82), (3.85), *and* (3.86).

### 3.3.2 With consumption

Based on the results quoted in the previous subsection, we are able to analyze the behaviors of the free boundaries when consumption applies. The following theorem makes it possible to extend most of the results in Proposition 5 to the consumption case.

**Lemma 3.3.2.** *Let $x_s(\tau)$ and $x_b(\tau)$ be the two free boundaries from problem* (3.11) *(optimal trading boundaries when consumption applies), and $\bar{x}_s(\tau)$ and $\bar{x}_b(\tau)$ be the two free boundaries from problem* (3.75) *(optimal trading boundaries in the absence of consumption). Then the following comparison results hold:*

$$x_s(\tau) \geq \bar{x}_s(\tau), \tag{3.90}$$

$$x_b(\tau) \geq \bar{x}_b(\tau). \tag{3.91}$$

The above proposition implies reasonable financial intuition that to maintain consumption, whose withdrawal is from the bank account, the investor prefers to keep a larger fraction of wealth in the bank account. Thus compared to the no consumption case, the no-transaction region shifts rightward in the $(x, \tau)$ plane.

We now prove it in terms of the maximum principle.

*Proof.* Let $v(x,\tau)$ and $\bar{v}(x,\tau)$ be the solution to problem (3.11) and (3.75), respectively. By (3.73), it follows that

$$\mathcal{L}_w v \leq 0.$$

Thanks to the maximum principle (cf. Friedman (1982), Page 74), we have

$$v(x,\tau) \geq \bar{v}(x,\tau).$$

It will lead to the inequalities that

$$\begin{aligned} \bar{v}(x,\tau) &< \frac{1}{x+1-\alpha}, \text{ if } v(x,\tau) < \frac{1}{x+1-\alpha}, \\ v(x,\tau) &> \frac{1}{x+1+\theta}, \text{ if } \bar{v}(x,\tau) > \frac{1}{x+1+\theta}, \end{aligned}$$

which directly yield (3.90) and (3.91). $\square$

**Remark 3.3.3.** *By* (3.81) *and* (3.90), *we infer* $x_s(\tau) \geq \bar{x}_s(\tau) > \lim_{\tau\to+\infty} \bar{x}_s(\tau) > -(1-\alpha)$. *Then we can choose* $x^* = \lim_{\tau\to+\infty} \bar{x}_s(\tau)$ *such that* $x_s(\tau)$ *never hits the line* $x = x^*$.

Based on Proposition 3.3.2, we give the following theorem on the behaviors of trading boundaries with consumption in presence:

**Theorem 3.3.4.** *Let* $x_s(\tau)$ *and* $x_b(\tau)$ *be the two free boundaries from problem* (3.11) *(the optimal trading boundaries with consumption), and let* $\tau_0$ *be as defined in* (3.79). *Then* $x_s(\tau) < x_b(\tau)$, *and*

*(i) for any* $\tau > 0$,

$$x_s(\tau) \leq x_s(0^+) \equiv \lim_{\tau\to\infty} x_s(\tau) = (1-\alpha)\, x_M; \tag{3.92}$$

*moreover*

$$x_s(\tau) \equiv 0 \quad \text{if } \mu - r - \gamma\sigma^2 = 0 \tag{3.93}$$

$$x_s(\tau) < 0 \quad \text{if } \mu - r - \gamma\sigma^2 > 0 \tag{3.94}$$

$$x_s(\tau) > 0 \quad \text{if } \mu - r - \gamma\sigma^2 < 0 \tag{3.95}$$

*(ii) for any* $\tau > 0$,

$$x_b(\tau) \geq (1+\theta)\, x_M, \tag{3.96}$$

*and*

$$x_b(\tau) = +\infty \quad \text{if and only if} \quad \tau \in (0, \tau_0]; \tag{3.97}$$

*Proof.* First, $x_s(\tau) < x_b(\tau)$ is clear. Second, (3.92) has been proved in Lemma 3.

If $\mu - r - \gamma\sigma^2 > 0$, then $x_M < 0$ and (3.94) is a direct conclusion from (3.92).

When $\mu - r - \gamma\sigma^2 < 0$, (3.95) follows from (3.83) and (3.90).

When $\mu - r - \gamma\sigma^2 = 0$, we again can use (3.83) and (3.90) to get $x_s(\tau) \geq 0$, while by (3.92), we have $x_s(\tau) \leq 0$. Thus the only possibility should be $x_s(\tau) \equiv 0$.

(3.96) is a direct conclusion from the combination of (3.85) and (3.91). Or, it can be similarly proved as for (3.92), i.e. for any $(x,t) \in \mathbf{BR}$, it should hold that

$$0 \leq \left(\frac{\partial}{\partial\tau} - \mathcal{L} + \mathcal{L}_w\right)\left(\frac{1}{x+1+\theta}\right) = \frac{1+\theta}{(x+1+\theta)^3}\left[(\mu - r)\, x + (1+\theta)\left(\mu - r - \gamma\sigma^2\right)\right],$$

so it must be true that $x_b(\tau) \geq (1+\theta)\, x_M$.

The last remains to prove is (3.97). It is obvious from (3.86) and (3.91) that when $\tau \in (0, \tau_0]$, $x_b(\tau) = +\infty$.

Then we prove the other direction, say, when $\tau > \tau_0$, $x_b(\tau) < +\infty$.

By transformation

$$\begin{aligned} &z = \tfrac{x}{x+1+\lambda}, \\ &\widetilde{v}(z,\tau) = \left(v(x,\tau) - \tfrac{1}{x+1+\lambda}\right) \tfrac{(x+1+\lambda)^2}{1+\lambda}, \end{aligned}$$

problem (3.11) becomes

$$\begin{cases} \widetilde{v}_\tau - \widetilde{\mathcal{L}}\widetilde{v} + \widetilde{\mathcal{L}}_w \widetilde{v} = 0, & \text{if } 0 < \widetilde{v} < \frac{\theta+\alpha}{(1-\alpha)+(\theta+\alpha)z}, \\ \widetilde{v}_\tau - \widetilde{\mathcal{L}}\widetilde{v} + \widetilde{\mathcal{L}}_w \widetilde{v} \geq 0, & \text{if } \widetilde{v} = 0, \\ \widetilde{v}_\tau - \widetilde{\mathcal{L}}\widetilde{v} + \widetilde{\mathcal{L}}_w \widetilde{v} \leq 0, & \text{if } \widetilde{v} = \frac{\theta+\alpha}{(1-\alpha)+(\theta+\alpha)z}, \\ \widetilde{v}(z,0) = \frac{\theta+\alpha}{(1-\alpha)+(\theta+\alpha)z}. \end{cases} \tag{3.98}$$

in $\frac{1-\alpha}{\theta+\alpha} < z < 1$, $\tau > 0$. Here

$$\begin{aligned} \widetilde{\mathcal{L}}\widetilde{v} \;=\; & \frac{1}{2}\sigma^2 z^2 (1-z)^2 \widetilde{v}_{zz} - \left((\mu - r - (1+\gamma)\sigma^2) + 3\sigma^2 z\right) z(1-z)\widetilde{v}_z \\ & - \left(\mu - r - \gamma\sigma^2 - 2(\mu - r - (1+\gamma)\sigma^2)z - 3\sigma^2 z^2\right)\widetilde{v} \\ & - \left(\sigma^2 z + \mu - r - \gamma\sigma^2\right) + (1-\gamma)\,\sigma^2 z[1 + (1-z)\widetilde{v}]\,[z(1-z)\widetilde{v}_z + (1-2z)\widetilde{v} + 1] \end{aligned}$$

and

$$\widetilde{\mathcal{L}}_w \widetilde{v} = \left(e^{(1-\gamma)w}\,[1 + (1-z)\,\widetilde{v}]\,\frac{1-z}{1+\theta}\right)^{-\frac{1}{\gamma}} \frac{(1-z)^2}{1+\theta}\left(\widetilde{v}^2 + \widetilde{v}_z\right).$$

Define

$$z_b(\tau) = \sup_z \left\{ z \in \left(\frac{1-\alpha}{\theta+\alpha}, 1\right) : \tilde{v}\,(z,\tau) > 0 \right\}$$

Clearly

$$z_b(\tau) = \frac{x_b(\tau)}{x_b(\tau) + 1 + \theta}. \tag{3.99}$$

To prove $x_b(\tau) < +\infty$, if $\tau > \tau_0$, it suffices to show $z_b(\tau) < 1$, if $\tau > \tau_0$.

Noticing that $e^{(1-\gamma)w} = (1-\gamma)\,V$, it is not hard to verify that $\widetilde{\mathcal{L}}_w \widetilde{v}|_{z=1} = 0$. Therefore, at $z = 1$, problem (3.98) is reduced to

$$\begin{cases} \widetilde{v}_\tau(1,\tau) - (\mu - r)\,\widetilde{v}(1,\tau) + \mu - r = 0 & \text{if } \widetilde{v}(1,\tau) > 0, \\ \widetilde{v}_\tau(1,\tau) - (\mu - r)\,\widetilde{v}(1,\tau) + \mu - r \geq 0 & \text{if } \widetilde{v}(1,\tau) = 0, \quad \tau > 0, \\ \widetilde{v}(1,0) = \frac{\theta+\alpha}{1+\theta}, & \end{cases}$$

whose solution is

$$\widetilde{v}(1,\tau) = \max\left(1 - e^{(\mu-r)\tau}\frac{1-\alpha}{1+\theta}, 0\right) = \begin{cases} 1 - e^{(\mu-r)\tau}\frac{1-\alpha}{1+\theta} \text{ when } \tau \in (0,\tau_0], \\ 0 \text{ when } \tau > \tau_0. \end{cases}$$

which implies the desired result. The proof is complete. □

We complete this chapter by the following 3 remarks:

**Remark 3.3.5.** *Compared with the no-consumption case, the monotonicity of free boundaries is not available, since it does not hold that $v_\tau \leq 0, \quad \forall \tau > 0$ (note that it still holds at $\tau = 0$). A numerical example about the non-monotonicity is presented in Dai and Zhong (2008). In addition,* (3.87)-(3.89) *means that in the no-consumption case, $x_b(\tau)$ intersects with $x = 0$ at $\tau_1$. However, when consumption is present, this property is no longer true due to the additional term $\mathcal{L}_w v$ caused by consumption. All theoretical results in this section are numerically demonstrated by Dai and Zhong (2008).*

**Remark 3.3.6.** *In the case when transaction costs are absent, Merton has shown that an investor should never leverage if the risk premium $\mu - r - \gamma\sigma^2$ is non-positive. This remains true when transaction costs apply. Indeed, from* (3.93)-(3.95)*, we infer that $x_b \geq x_s \geq 0$ if and only if when $\mu - r - \gamma\sigma^2 \leq 0$, which implies the conclusion.*

**Remark 3.3.7.** (3.97) *indicates that there is a critical time after which it is never optimal to purchase stocks. This is one interesting and important feature of the finite horizon problem. It's counterpart* (3.86) *in the no-consumption case was first found by Liu and Loewenstein (2002). The intuition behind this is that if the investor does not have an expected horizon long enough to recover at least the transaction costs, then s/he should not purchase any additional stock.*

Chapter 4

# Market closure, portfolio selection and liquidity premium

Most literature on portfolio selection problem considers a continuously opening market and trading is allowed at any time. However, in reality, periodic closure during nights, weekends, public holidays do apply in almost all financial markets. In this chapter, we will investigate how market closure will affect portfolio selection in the absence of consumption.

## 4.1 The model

Following the idea in previous chapter, we consider an investor with CRRA utility function

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma} - \frac{1}{1-\gamma},$$

where $\gamma > 0$ is the constant relative risk aversion coefficient. To simplify the model, we assume there is no consumption. In this case, the investor's objective is to maximize his expected utility only from terminal liquidation wealth at some finite horizon $T \in (0, \infty)$.

Different from the standard literature, we assume that the stock market closes periodically. Specifically, the investment horizon is partitioned into $0 = t_0 < ... < ... <$

$t_{2N+1} = T$. Market is open ("daytime") in time intervals $[t_{2i}, t_{2i+1}]$ and trading is allowed; while the market is closed ("nighttime") in $(t_{2i+1}, t_{2i+2}), \forall i = 0, 1, ..., N$. Of course these intervals can be of different length, and thus can deal with closure on weekends and holidays.

As in previous chapter, we assume proportional transaction costs prevail in market. So when market is open, the investor can buy the stock at the ask price $S_t^A = (1+\theta)S_t$ and sell the stock at the bid price $S_t^B = (1-\alpha)S_t$, where $\theta \geq 0$ and $0 \leq \alpha < 1$ represent the proportional transaction costs.

Again the investor can invest in two financial instruments: a risk-less asset (bank account) and a risky asset (stock). The bond price process $P_t$ evolves regularly as in standard literature, which is

$$dP_t = rP_t dt.$$

The stock price process $S_t$, however, in the presence of market closure, is governed by an SDE with periodic drift and diffusion parameters:

$$dS_t = \mu(t) S_t dt + \sigma(t) S_t dB_t,$$

with

$$\mu(t) = \begin{cases} \mu_d, & \text{day} \\ \mu_n, & \text{night} \end{cases} \quad \text{and} \quad \sigma(t) = \begin{cases} \sigma_d, & \text{day} \\ \sigma_n, & \text{night,} \end{cases}$$

where $\mu_d > r, \mu_n > r, \sigma_d > 0, \sigma_n > 0$ are all constants and $\{B_t; t \geq 0\}$ is a one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ with $B_0 = 0$ almost surely. We assume $\mathcal{F} = \mathcal{F}_\infty$, the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous and each $\mathcal{F}_t$ contains all null sets of $\mathcal{F}_\infty$.

Assume that a CRRA investor holds $X_t$ and $Y_t$ in bank and stock respectively, the counterparts of (3.1) and (3.2) are now as follows:

$$dX_t = rX_t dt - (1+\theta)dI_t + (1-\alpha)dD_t \tag{4.1}$$

$$dY_t = \mu(t) Y_t dt + \sigma(t) Y_t dB_t + dI_t - dD_t, \tag{4.2}$$

where $I_t$ and $D_t$ are the cumulative stock sales and purchases processes, which are right-continuous (with left hand limits), nonnegative, and nondecreasing $\{\mathcal{F}_t\}_{t\geq 0}$-adapted processes with $I_0 = D_0 = 0$. As a contrast to the case where market is continuously open, trading is not allowed during night, so it holds that

$$dI_t = 0,\ dD_t = 0,\ \text{if } t \in (t_{2i+1}, t_{2i+2}).$$

Again the time-$t$ wealth after liquidation of the investor is

$$\begin{aligned} W_t &= \begin{cases} X_t + (1-\alpha)\,Y_t, & \text{if } Y_t \geq 0, \\ X_t - (1-\theta)\,Y_t, & \text{if } Y_t < 0, \end{cases} \\ &= X_t + (1-\alpha)\,Y_t^+ - (1-\theta)\,Y_t^-. \end{aligned}$$

It is required that the investor's net wealth should be positive, i.e.,

$$W_t \geq 0,\ \forall t \geq 0, \tag{4.3}$$

Because the investor cannot trade when market is closed and the stock price can get arbitrarily close to 0 and is unbounded above, the solvency constraint (4.3) implies that the investor cannot borrow or shortsell at market close. So the solvency region is different between trading and non-trading periods, i.e.,

$$\mathcal{S} = \begin{cases} \mathcal{S}_d = \left\{(x,y) \in \mathbb{R}^2 : x + (1+\theta)y > 0,\ x + (1-\alpha)y > 0\right\}, & \text{if } t \in [t_{2i}, t_{2i+1}], \\ \mathcal{S}_n = \left\{(x,y) \in \mathbb{R}^2 : x > 0,\ y > 0\right\}, & \text{if } t \in (t_{2i+1}, t_{2i+2}). \end{cases}$$

A trading strategy $(I_t, D_t)$ is admissible for $(x,y)$ starting from some time $s \in [0,T]$, if $(X_t, Y_t)$ with $X_s = x,\ Y_s = y$ evolves within $\mathcal{S}$ for all $t \in [s,T]$. Let $\mathcal{A}_s(x,y)$ denote the set of all admissible trading strategies such that the investor is always solvent (4.3) under the governance of (4.1) and (4.2). The investor's problem is then to maximize the expected utility from terminal wealth over all admissible trading strategies, or in a mathematical word, to look for

$$\sup_{(I,D)\in\mathcal{A}_0(x,y)} E\left[U(W_T)\right], \tag{4.4}$$

where $x,\ y$ are some given initial positions in the bank account and stock respectively.

## 4.2 Optimal strategy without transaction costs

For the purpose of comparison, let us first consider the case without transaction costs (i.e., $\alpha = \theta = 0$).

Since the investor can only trade during in the "daytime" when market is open, his control problem is only definable at the time $t \in [t_{2i}, t_{2i+1}]$. In this case, let $W_s = X_s + Y_s$ be the time $s$ wealth for $s \geq t$. Then the investor's problem at time $t$ becomes to maximize his expected utility from terminal wealth over all admissible "daytime" strategies.

When trading is allowed, as in Merton's model where transaction costs are absent, we are able to reduce the dimension of problem by introducing control $\pi(t)$ instead of $(I,\ D)$, where $\pi(t)$ represents the fraction of wealth invested in the stock at time $s$. Then we define the time-$t$ ($t \in [t_{2i}, t_{2i+1}]$) value function by

$$V(w,t) \equiv \sup_{\{\pi\} \in \mathcal{A}_t(w)} E_t^w \left[ U(W_T) \right], \quad t \in [t_{2i}, t_{2i+1}], \tag{4.5}$$

subject to the solvency constraint (4.3) and the self financing condition

$$\begin{aligned} dW_s &= rW_s ds + \pi(s) W_s (\mu_d - r) ds + \pi(s) W_s \sigma_d dB_s, \quad \forall s \geq t, \\ W_t &= w. \end{aligned}$$

Note that trading is not allowed during market closure. Thus with the bank account value process and the stock value process evolving independently, it would not be possible to reduce the dimension of problem by introducing $\pi(t)$ any more. So the "nighttime" value function should be of the form $V(x, y, t), \quad t \in (t_{2i+1}, t_{2i+2})$.

The basic idea for this investor's problem is to solve backward iteratively from the last trading period ("daytime"), then the last non-trading period ("nighttime"), and then the second last daytime, the second last nighttime, so on and so forth.

Before we present the solution to problem (4.4), we need first look at three sub problems, which respectively correspond to the optimization during trading time, the evolution during night time, and the adjustment at market close.

### 4.2.1 Three subproblems

Before moving to the optimization problem over an arbitrary interval $[t,\, T)$, we consider the following three problems within one open-close period

$$[t_{2i-1},\ t_{2i+1}) = \{t_{2i-1}\} \cup (t_{2i-1},\ t_{2i}) \cup [t_{2i},\ t_{2i+1})$$

first:

- **Utility optimization within one trading period**

For $t \in [t_{2i},\, t_{2i+1})\,,\ i = 0, 1,\, \ldots,\, N$ , market is open such that one can optimize his portfolio via the control variable $\pi$. We define the value function during "daytime" as follows:

$$J_d^i\,(x,\, y,\, t) \equiv \sup_{\{\pi(s):\, s\in[t,\, t_{2i+1})\}} E_t^{x,y}\left[\left(X_{t_{2i+1}} + Y_{t_{2i+1}}\right)^{1-\gamma}\right] \tag{4.6}$$

- **Utility evolution within one non-trading period**

For $t \in (t_{2i-1},\, t_{2i})\,,\ i = 0, 1,\, \ldots,\, N$, market is closed and the control problem degenerates to an expectation function:

$$J_n^i\,(x,\, y,\, t) \equiv E_t^{x,y}\left[(X_{t_{2i}} + Y_{t_{2i}})^{1-\gamma}\right] \tag{4.7}$$

- **Utility optimization at when market closes**

At $t = t_{2i-1},\ i = 1, 2,\, \ldots,\, N$ , market closes. One should trade instantaneously to optimize his expected utility at the moment when market opens next time, say: $t = t_{2i}$. We define this subproblem as

$$J_n^i\,(x,\, y,\, t_{2i-1})^* \equiv \sup_{\{\pi(t_{2i-1})\in[0,\, 1]\}} E_{t_{2i-1}}^{x,\, y}\left[(X_{t_{2i}} + Y_{t_{2i}})^{1-\gamma}\right] \tag{4.8}$$

Here we require $\pi\,(t_{2i-1}) \in [0,\, 1]$ to exclude leverage during market closure and thus to diminish the possibility of bankruptcy.

For problem (4.6)-(4.8) defined above, the following propositions hold:

**Proposition 6.** *Let $J_d^i(x,\, y,\, t)$ be the value function defined in equation* (4.6)*, Then*

$$J_d^i(x,\, y,\, t) = (x+y)^{1-\gamma} \cdot e^{(1-\gamma)\eta_d(t_{2i+1}-t)} \tag{4.9}$$

*where*

$$\eta_d = r + \frac{(\mu_d - r)^2}{2\gamma\sigma_d^2}; \tag{4.10}$$

*and $J_d^i(x,\, y,\, t)$ is achieved by the following optimal trading strategy:*

$$\pi(s)^* = \frac{\mu_d - r}{\gamma\sigma_d^2} = \pi_M, \quad s \in [t,\, t_{2i+1}).$$

*Proof.* This is actually the well-known Merton's result, where the only difference lies in terminal condition. However, the different terminal condition will not affect the proof, which we present briefly below:

For $t \in [t_{2i},\, t_{2i+1})$, we rewrite $J_d^i(x,\, y,\, t)$ as $J_d^i(w,\, t)$ where $w = x+y$. Since trading is allowed during daytime, the stochastic control problem (4.6) leads to the following HJB equation:

$$\sup_{\{\pi_t\}} \left\{ \left(J_d^i\right)_t + \frac{1}{2}\pi_t^2\sigma_d^2 w^2 \left(J_d^i\right)_{ww} + \left[r + \pi_t(\mu_d - r)\right] w \left(J_d^i\right)_w \right\} = 0; \tag{4.11}$$

with the terminal condition

$$J_d^i(w,\, t_{2i+1}) = w^{1-\gamma}$$

By considering first order derivative, the maximum is attained at

$$\pi(t)^* = -\frac{(\mu_d - r)\left(J_d^i\right)_w}{\sigma_d^2 w \left(J_d^i\right)_{ww}},$$

in this case, equation (4.11) has a closed form solution

$$J_d^i(w,\, t) = w^{1-\gamma} \cdot e^{(1-\gamma)\eta_d(t_{2i+1}-t)}$$

or equivalently

$$J_d^i(x,\, y,\, t) = (x+y)^{1-\gamma} \cdot e^{(1-\gamma)\eta_d(t_{2i+1}-t)}$$

and the optimal trading strategy is given by

$$\pi(t)^* = \frac{\mu_d - r}{\gamma\sigma_d^2} = \pi_M$$

□

**Proposition 7.** *Let* $J_n^i(x,\, y,\, t)$ *be the value function defined in equation* (4.7)*, Then*

$$J_n^i(x,\, y,\, t) = (x+y)^{1-\gamma} \cdot e^{(1-\gamma)r(t_{2i}-t)} \cdot G_i\left(\frac{y}{x+y},\, t\right) \tag{4.12}$$

*where*

$$G_i(\pi,\, t) = E\left\{[1 + \pi\,(R\,(t_{2i}, t) - 1)]^{1-\gamma}\right\} \tag{4.13}$$

*and*

$$R\,(u, v) = \exp\left[\left(\mu_n - r - \frac{\sigma^2}{2}\right)(u - v) + \sigma_n\,(B_u - B_v)\right]$$

*Proof.* For $t \in (t_{2i-1},\, t_{2i})$, since no control is applicable, the SDEs governing $X_s$ and $Y_s$ degenerate to

$$\begin{cases} dX_s = rX_s ds \\ dY_s = \mu_n Y_s ds + \sigma_n Y_s dB_s \end{cases} \quad \forall s \geq t; \tag{4.14}$$

Given

$$(X_t,\, Y_t) = (x,\, y)$$

SDEs (4.14) have the following solution

$$\begin{cases} X_s = x \cdot \exp\left[r\,(s - t)\right] \\ Y_s = y \cdot \exp\left[\left(\mu_n - \frac{\sigma_n^2}{2}\right)(s - t) + \sigma_n\,(B_s - B_t)\right] \end{cases}$$

So

$$\begin{aligned} J_n^i(x,\, y,\, t) &= E\left\{\left[x \cdot e^{r(t_{2i}-t)} + y \cdot e^{\left(\mu_n - \frac{\sigma_n^2}{2}\right)(t_{2i}-t) + \sigma_n\left(B_{t_{2i}} - B_t\right)}\right]^{1-\gamma}\right\} \\ &= (x+y)^{1-\gamma}\, e^{(1-\gamma)r(t_{2i}-t)} E\left\{\left[1 + \frac{y}{x+y}\,(R\,(t_{2i},\, t) - 1)\right]^{1-\gamma}\right\} \\ &\equiv (x+y)^{1-\gamma}\, e^{(1-\gamma)r(t_{2i}-t)} G_i\left(\frac{y}{x+y},\, t\right) \end{aligned}$$

□

**Proposition 8.** *Let* $J_n^i(x,\, y,\, t_{2i-1})^*$ *be the value function defined in equation* (4.8)*, Then*

$$J_n^i(x,\, y,\, t_{2i-1})^* = (x+y)^{1-\gamma} \cdot e^{(1-\gamma)r(t_{2i} - t_{2i-1})} \cdot G_i^*$$

*where*

$$G_i^* \equiv E\left\{\left[1+\pi_i^*\left(R_i-1\right)\right]^{1-\gamma}\right\} \tag{4.15}$$

*and* $\pi_i^*$ *solves the optimization problem:*

$$\sup_{\pi(t_{2i-1})\in[0,\,1]} E\left\{\left[1+\pi\left(t_{2i-1}\right)\cdot\left(R_i-1\right)\right]^{1-\gamma}\right\} \tag{4.16}$$

*with*

$$R_i = R(t_{2i}, t_{2i-1}) = \exp\left[\left(\mu_n - r - \sigma_n^2/2\right)\left(t_{2i}-t_{2i-1}\right) + \sigma_n(B(t_{2i}) - B(t_{2i-1}))\right]$$

*Proof.* For a position $\left(X_{t_{2i-1}}, Y_{t_{2i-1}}\right) = (x,\, y)$, if no control is allowed at time $t_{2i-1}$, the value function would be the same as in (4.12) by putting $t = t_{2i-1}$. With control available, one can trade and adjust his position to $\pi_i^*$, which optimize $G_i\left(\frac{y}{x+y},\, t_{2i-1}\right)$ to $G_i^*$. □

**Remark 4.2.1.** *If* $\mu_n - r - \gamma\sigma_n^2 \leq 0$*, then one can find* $\pi_i^* \in [0,\,1]$ *which solves the first order condition equation of* (4.16)*:*

$$E\left\{\left[1+\pi_i^*\left(R_i-1\right)\right]^{-\gamma}\cdot\left(R_i-1\right)\right\} = 0 \tag{4.17}$$

*However, if* $\mu_n - r - \gamma\sigma_n^2 > 0$*, equation* (4.17) *does not have a root in* $[0,\,1]$*; in this case, one takes* $\pi_i^* = 1$.

### 4.2.2 Value function with market closure in the absence of transaction costs

We summarize the main result for this case of no transaction costs in the following theorem (with the convention that $t_{-1} = 0$).

**Theorem 4.2.2.** *Suppose that* $\alpha = \theta = 0$*. Then the value function for* $t \in [t_{2i-1},\, t_{2i+1})$, $i = 0,\, 1,\, \ldots,\, N$ *is given by*

$$V\left(x,\,y,\,t\right) = \begin{cases} \frac{(x+y)^{1-\gamma}}{1-\gamma} e^{(1-\gamma)\eta(t)}\left(\prod_{k=i+1}^{N} G_k^*\right) - \frac{1}{1-\gamma}, & t \in [t_{2i},\, t_{2i+1})\,; \\ \frac{(x+y)^{1-\gamma}}{1-\gamma} e^{(1-\gamma)\eta(t)}\left(\prod_{k=i+1}^{N} G_k^*\right) G_i\left(\frac{y}{x+y},\, t\right) - \frac{1}{1-\gamma}, & t \in (t_{2i-1},\, t_{2i}) \end{cases} \tag{4.18}$$

*and it is attained by the optimal trading policy of*

$$\pi(t)^* = \begin{cases} \pi_M, & t \in [t_{2i}, t_{2i+1}); \\ \frac{\pi_i^* R(t, t_{2i-1})}{1+\pi_i^*(R(t, t_{2i-1})-1)}, & t \in [t_{2i-1}, t_{2i}), \end{cases}$$

*where*

$$G_i(\pi, t) = E\left\{[1 + \pi(R(t_{2i}, t) - 1)]^{1-\gamma}\right\}, \tag{4.19}$$

$$R(u, v) = \exp\left[\left(\mu_n - r - \sigma_n^2/2\right)(u - v) + \sigma_n(B_u - B_v)\right], \tag{4.20}$$

$$\pi_i^* = \arg\max_{\pi \in [0,1]} G_i(\pi, t_{2i-1}), \quad G_i^* = G_i(\pi_i^*, t_{2i-1}), \tag{4.21}$$

*and*

$$\eta(t) = r(T - t) + \frac{(\mu_d - r)^2}{2\gamma\sigma_d^2} \sum_{i=0}^{N} (t_{2i+1} - t_{2i} \vee t)^+ \tag{4.22}$$

*Proof.* We complete the proof by a mathematical induction backward in time. In the argument, we are going to use the following notation

$$\begin{aligned} L_d(t) &\equiv \sum_{k=0}^{N} (t_{2k+1} - t_{2k} \vee t)^+ \\ &= \begin{cases} \sum_{k=i+1}^{N} (t_{2k+1} - t_{2k}) + t_{2i+1} - t, & \text{if } t \in [t_{2i}, t_{2i+1}); \\ \sum_{k=i}^{N} (t_{2k+1} - t_{2k}), & \text{if } t \in [t_{2i-1}, t_{2i}). \end{cases} \end{aligned} \tag{4.23}$$

It stands for the cumulative time of sub-intervals of $[t, T)$ during which market is open. And note that

$$\begin{aligned} \eta(t) &\equiv \eta_d L_d(t) + r(T - t - L_d(t)) \\ &= r(T - t) + \frac{(\mu_d - r)^2}{2\gamma\sigma_d^2} L_d(t) \end{aligned} \tag{4.24}$$

1. First we prove that when $i = N$, for $t \in [t_{2N-1}, t_{2N+1})$, (4.18) is true.

   (a) For $t \in [t_{2N}, t_{2N+1})$, the value function is

$$\begin{aligned} V(x, y, t) &= \sup_{\{\pi(s):\, s \in [t, t_{2N+1})\}} E_t^{x,y}\left\{\frac{1}{1-\gamma}\left[(X_T + Y_T)^{1-\gamma} - 1\right]\right\} \\ &= \frac{1}{1-\gamma}\left\{\sup_{\{\pi(s):\, s \in [t, t_{2N+1})\}} E_t^{x,y}\left[\left(X_{t_{2N+1}} + Y_{t_{2N+1}}\right)^{1-\gamma}\right] - 1\right\} \\ &= \frac{1}{1-\gamma}\left\{J_d^N(x, y, t) - 1\right\} \end{aligned}$$

By Proposition 6,

$$V\left(x,\,y,\,t\right)=\frac{1}{1-\gamma}\left\{\left(x+y\right)^{1-\gamma}\cdot e^{(1-\gamma)\eta(t)}-1\right\}$$

(b) For $t\in\left(t_{2N-1},\,t_{2N}\right)$, the value function $V\left(x,\,y,\,t\right)$ is

$$\begin{aligned}
& \sup_{\{\pi(s):\,s\in[t_{2N},\,t_{2N+1})\}} E_t^{x,\,y}\left\{\frac{1}{1-\gamma}\left[\left(X_T+Y_T\right)^{1-\gamma}-1\right]\right\} \\
= \;& \frac{1}{1-\gamma}\left\{E_t^{x,\,y}\left(\sup_{\{\pi(s):\,s\in[t_{2N},\,t_{2N+1})\}} E_{t_{2N}}^{X_{2N},\,Y_{t_{2N}}}\left[\left(X_{t_{2N+1}}+Y_{t_{2N+1}}\right)^{1-\gamma}\right]\right)-1\right\} \\
= \;& \frac{1}{1-\gamma}\left\{E_t^{x,\,y}\left[J_d^N\left(X_{t_{2N}},\,Y_{t_{2N}},t_{2N}\right)\right]-1\right\} \\
= \;& \frac{1}{1-\gamma}\left\{E_t^{x,\,y}\left[\left(X_{t_{2N}}+Y_{t_{2N}}\right)^{1-\gamma}\right]\cdot e^{(1-\gamma)\eta_d(t_{2N+1}-t_{2N})}-1\right\} \\
= \;& \frac{1}{1-\gamma}\left\{J_n^N\left(x,\,y,\,t\right)\cdot e^{(1-\gamma)\eta_d L_d(t)}-1\right\}
\end{aligned}$$

By Lemma 7,

$$V\left(x,\,y,\,t\right)\;=\;\frac{1}{1-\gamma}\left\{\left(x+y\right)^{1-\gamma}G_N\left(\frac{y}{x+y},\,t\right)\cdot e^{(1-\gamma)\eta(t)}-1\right\}.$$

(c) When $t=t_{2N-1}$, define the control set

$$C^N\left(\pi\right)\equiv\left\{\pi\left(s\right):\;s\in\left\{t_{2N-1}\cup\left[t_{2N},\,t_{2N+1}\right)\right\};\,\pi\left(t_{2N-1}\right)\in\left[0,\,1\right]\right\},$$

and the value function is

$$\begin{aligned}
& V\left(x,\,y,\,t_{2N-1}\right) \\
= \;& \sup_{C^N(\pi)} E_{t_{2N-1}}^{x,\,y}\left\{\frac{1}{1-\gamma}\left[\left(X_T+Y_T\right)^{1-\gamma}-1\right]\right\} \\
= \;& \left\{\sup_{\pi(t_{2N-1})} E_{t_{2N-1}}^{x,\,y}\left(\sup_{\pi(s)} E_{t_{2N}}^{X_{2N},\,Y_{t_{2N}}}\left[\left(X_{t_{2N+1}}+Y_{t_{2N+1}}\right)^{1-\gamma}\right]\right)-1\right\}\Big/\left(1-\gamma\right) \\
= \;& \frac{1}{1-\gamma}\left\{\sup_{\pi(t_{2N-1})} E_{t_{2N-1}}^{x,\,y}\left(J_d^N\left(X_{t_{2N}},\,Y_{t_{2N}},t_{2N}\right)\right)-1\right\} \\
= \;& \frac{1}{1-\gamma}\left\{\sup_{\pi(t_{2N-1})} E_{t_{2N-1}}^{x,\,y}\left[\left(X_{t_{2N}}+Y_{t_{2N}}\right)^{1-\gamma}\right]\cdot e^{(1-\gamma)\eta_d(t_{2N+1}-t_{2N})}-1\right\} \\
= \;& \frac{1}{1-\gamma}\left[J_n^N\left(x,\,y,\,t_{2N-1}\right)^*\cdot e^{(1-\gamma)\eta_d L_d(t)}-1\right].
\end{aligned}$$

By Corollary 8, the value function is

$$\begin{aligned} J(x,\, y,\, t_{2N-1}) &= \frac{1}{1-\gamma}\left[(x+y)^{1-\gamma}\, e^{(1-\gamma)r(t_{2N}-\, t_{2N-1})} G_N^* \cdot e^{(1-\gamma)\eta_d L_d(t)} - 1\right] \\ &= \frac{1}{1-\gamma}\left[(x+y)^{1-\gamma}\, e^{(1-\gamma)\eta(t)} G_N^* - 1\right]. \end{aligned}$$

2. Next we suppose (4.18) is true for $i = k+1$ or alternatively speaking, $t \in [t_{2k+1},\, t_{2k+3})$, we hope to provle that (4.18) is also true for $i = k$, $t \in [t_{2k-1},\, t_{2k+1})$

   (a) For $t \in [t_{2k},\, t_{2k+1})$

$$\begin{aligned} & J(x,\, y,\, t) \\ = & \sup_{\pi(s):\, s\in[t,\, t_{2N+1})} E_t^{x,\, y}\left\{\frac{1}{1-\gamma}\left[(X_T+Y_T)^{1-\gamma}-1\right]\right\} \\ = & \sup_{\pi(s):\, s\in[t,\, t_{2k+1})} E_t^{x,\, y} \\ & \left\{\sup_{\pi(s):\, s\in[t_{2k+1},\, t_{2N+1})} E_{t_{2k+1}}^{X_{t_{2k+1}},\, Y_{t_{2k+1}}}\left(\frac{1}{1-\gamma}\left[(X_T+Y_T)^{1-\gamma}-1\right]\right)\right\} \\ = & \sup_{\pi(s):\, s\in[t,\, t_{2k+1})} E_t^{x,\, y} \\ & \left\{\frac{1}{1-\gamma}\left[\left(X_{t_{2k+1}}+Y_{t_{2k+1}}\right)^{1-\gamma} e^{(1-\gamma)\eta(t_{2k+1})}\left(\prod_{j=k+1}^{N} G_j^*\right)-1\right]\right\} \\ = & \frac{1}{1-\gamma}\cdot \\ & \left\{\sup_{\pi(s):\, s\in[t,\, t_{2k+1})} E_t^{x,\, y}\left[\left(X_{t_{2k+1}}+Y_{t_{2k+1}}\right)^{1-\gamma}\right] e^{(1-\gamma)\eta(t_{2k+1})}\left(\prod_{j=k+1}^{N} G_j^*\right)-1\right\} \\ = & \frac{1}{1-\gamma}\left\{J_d^k(x,\, y,\, t)\, e^{(1-\gamma)\eta(t_{2k+1})}\left(\prod_{j=k+1}^{N} G_j^*\right)-1\right\} \\ = & \frac{1}{1-\gamma}\left\{(x+y)^{1-\gamma}\, e^{(1-\gamma)\eta(t)}\left(\prod_{j=k+1}^{N} G_j^*\right)-1\right\} \end{aligned}$$

(b) For $t \in (t_{2k-1},\, t_{2k})$, trading is not available and

$$
\begin{aligned}
& J\left(x,\, y,\, t\right) \\
= \;& \sup_{\pi(s):\, s\in[t,\, t_{2N+1})} E_t^{x,\, y}\left\{\frac{1}{1-\gamma}\left[\left(x_T+y_T\right)^{1-\gamma}-1\right]\right\} \\
= \;& E_t^{x,\, y}\left\{\sup_{\pi(s):\, s\in[t_{2k},\, t_{2N+1})} E_{t_{2k}}^{x_{t_{2k}},\, y_{t_{2k}}}\left(\frac{1}{1-\gamma}\left[\left(x_T+y_T\right)^{1-\gamma}-1\right]\right)\right\} \\
= \;& E_t^{x,\, y}\left\{J\left(x_{t_{2k}}\, y_{t_{2k}},\, t_{2k}\right)\right\} \\
= \;& \frac{1}{1-\gamma}\left\{E_t^{x,\, y}\left[\left(x_{t_{2k}}+y_{t_{2k}}\right)^{1-\gamma}\right] e^{(1-\gamma)\eta(t_{2k})}\left(\prod_{j=k+1}^{N} G_j^*\right)-1\right\} \\
= \;& \frac{1}{1-\gamma}\left\{\left(x+y\right)^{1-\gamma} e^{(1-\gamma)\eta(t)}\left(\prod_{j=k+1}^{N} G_j^*\right) G_k\left(x,\, y,\, t\right)-1\right\} \qquad (4.25)
\end{aligned}
$$

(c) For $t = t_{2k-1}$, trading is available and putting

$$G_k\left(x,\, y,\, t_{2k-1}\right) = G_k^*$$

in (4.25) will yield

$$J\left(x,\, y,\, t_{2k-1}\right) = \frac{1}{1-\gamma}\left\{\left(x+y\right)^{1-\gamma} e^{(1-\gamma)\eta(t_{2k-1})}\left(\prod_{j=k}^{N} G_j^*\right) G_j^*-1\right\}$$

By the arguments in 1.(a.)-2.(c.), we have proved (4.18). □

The basic idea of solving for the optimal trading strategy is to solve the investor's problem period by period from time $T$. Our formulation allows arbitrary length of market open and closure. It is essential in the above proof that the value function always takes the form of

$$\frac{1}{1-\gamma}\left[\left(x+y\right)^{1-\gamma} A\left(t\right)-1\right],$$

where $A\left(t\right)$ only depends on $t$. This allows us to use the Merton's strategy in the day time and to repeat the derivation during each period $[t_{2i-1}, t_{2i+1})$ for any $i$.

Theorem 4.2.2 suggests that when market is open, the investor invests the same fraction of wealth in stock as in the case with no market closure. Then the investor adjusts his position at market close to take into account the effect of market closure and different return dynamics during night. In addition, since the investor cannot trade overnight, the stock position just before market open can be suboptimal and therefore another discrete adjustment is also likely at market open. The adjustments at market close and open suggests that the trading volume at these times are higher than in the rest of the trading hours, predicting a U-shaped trading volume pattern across trading hours.

Also since the support of stock price is from 0 to $\infty$, the investor can never buy on margin or shortsell at market close, otherwise solvency cannot be guaranteed. Thus when leverage is optimal, the effect of market closure on the optimal trading strategy should be greater.

One more thing interesting for the market closure model without transaction costs is that the optimal trading strategy during trading period is independent of parameter values during non-trading period. We will show later that this is no longer true in the presence of transaction costs.

### 4.2.3 Some variations of the optimal investment model without transaction costs

Following the idea of solving backward in time period by period, we would be able to derive closed form solutions for some variations of Merton's model. For later use in numerical analysis, we list the time-0 formulas of these value functions here.

To simplify expressions, we suppose one "open-close" period in the market is just one day (without consideration of weekends or public holidays), with equal "daytime" length $\Delta_d$ and "nighttime" length $\Delta_n$ for every day. Moreover, we assume there is 250 trading days in one year. The derivation of these value functions are straightforward by previously used approaches, we just omit them for concision.

- **Value function in Market A**[1] (day-night trading with $(\sigma_d,\, \sigma_n)$)

$$V_A\left(x,\, y,\, 0\right) = \frac{1}{1-\gamma}\left(x+y\right)^{1-\gamma}\exp\left[N\left(1-\gamma\right)\left(\eta_d^*\Delta_d + r\Delta_n + \frac{1}{1-\gamma}\log a^*\right)\right] - \frac{1}{1-\gamma} \tag{4.26}$$

where

$$a^* = E\left\{\left[1+\pi^*\left(e^{\left(\mu-r-\frac{\sigma_n^2}{2}\right)\Delta_n+\sigma_n\sqrt{\Delta_n}B_1}-1\right)\right]^{1-\gamma}\right\},$$

and

$$\pi^* = \arg\max_{\pi\in[0,\,1]} E\left\{\left[1+\pi\left(e^{\left(\mu-r-\frac{\sigma_n^2}{2}\right)\Delta_n+\sigma_n\sqrt{\Delta_n}B_1}-1\right)\right]^{1-\gamma}\right\};$$

- **Value function in Market B** (continuous trading with $(\sigma_d,\, \sigma_n)$)

$$V_B\left(x,\, y,\, 0\right) = \frac{1}{1-\gamma}\left(x+y\right)^{1-\gamma}\exp\left[N\left(1-\gamma\right)\left(\eta_d^*\Delta_d+\eta_n^*\Delta_n\right)\right] - \frac{1}{1-\gamma} \tag{4.27}$$

where

$$\begin{aligned} \eta_i^* &= r+\frac{\mu-r}{2\gamma\sigma_i^2}, \quad i=d,\, n, \\ N &= 250\times T \ :\ \text{number of days in } [0,\, T]; \end{aligned}$$

- **Value function in Market A** (day-night trading with $(\sigma_d,\, \sigma_n)$) **using the optimal strategy** $\pi_C$ **of Market C** (continuous trading with $(\sigma)$)

$$V_A^{\pi_C}\left(x,\, y,\, 0\right) = \frac{1}{1-\gamma}\left(x+y\right)^{1-\gamma}\exp\left[N\left(1-\gamma\right)\left(\eta_d\Delta_d + r\Delta_n + \frac{1}{1-\gamma}\log a\right)\right] - \frac{1}{1-\gamma} \tag{4.28}$$

where

$$\begin{aligned} \eta_d &= r+\frac{\left(\mu-r\right)^2}{\gamma\sigma^2}\left[1-\frac{\sigma_d^2}{2\sigma^2}\right], \\ a &= E\left\{\left[1+\min\left(\frac{\mu-r}{\gamma\sigma^2},\, 1\right)\cdot\left(e^{\left(\mu-r-\frac{\sigma_n^2}{2}\right)\Delta_n+\sigma_n\sqrt{\Delta_n}B_1}-1\right)\right]^{1-\gamma}\right\}. \end{aligned}$$

[1] This is just the simplified version of value function (4.18) with notations $\Delta_d$ and $\Delta_n$.

## 4.3 The transaction cost case

### 4.3.1 The value function and connection conditions

In the case where $\alpha + \theta > 0$ , the optimal investment problem is considerably more complicated. We still denote the investor's value function by $V(x, y, t)$, which represents the investor's problem at time $t$ :

$$V(x, y, t) \equiv \sup_{(I,\, D) \in\, \mathcal{A}_t(x,y)} E_t^{x,\, y}\left[U(W_T)\right].$$

Under regularity conditions on the value function, for $i = 0, 1, 2, ..., N$, we have the following Hamilton-Jacobi-Bellman (HJB) equations during trading period:

$$\min(-V_t - \mathscr{L}V, -(1-\alpha)V_x + V_y, (1+\theta)V_x - V_y) = 0, \quad \forall t \in [t_{2i}, t_{2i+1}), \tag{4.29}$$

and for non-trading period:

$$V_t + \mathscr{L}V = 0, \quad \forall t \in (t_{2i-1}, t_{2i}), \tag{4.30}$$

furthermore, at any market close before $T$, the investor faces an optimization problem, which is to trade $\Delta$ of stock in dollar value to maximize his value function. Thus the connection condition at market closes should be

$$V(x, y, t_{2i+1}) = \max_{\Delta \in \mathcal{C}(x,y)} V(x - (1+\theta)\Delta^+ + (1-\alpha)\Delta^-, y + \Delta, t_{2i+1}^+), \tag{4.31}$$

with the terminal condition

$$V(x, y, T) = \frac{(x + (1-\alpha)y^+ - (1+\theta)y^-)^{1-\gamma}}{1-\gamma} - \frac{1}{1-\gamma}, \tag{4.32}$$

where

$$\mathscr{L}V = \frac{1}{2}\sigma(t)^2 y^2 V_{yy} + \mu(t) y V_y + r x V_x,$$

and $\Delta$ takes value in the admissible set $\mathcal{C}(x, y)$ which maintains nonnegative liquidation wealth of this investor:

$$\mathcal{C}(x, y) = \{\Delta \in \mathbb{R} : x - (1+\theta)\Delta^+ + (1-\alpha)\Delta^- \geq 0, y + \Delta \geq 0\}.$$

As we show later, (4.29) implies that the solvency region $\mathcal{S}_d$ at each point during a day splits into a "Buy" region (BR), a "No-transaction" region (NT), and a "Sell" region (SR), as in Davis and Norman (1990).

The following verification theorem shows the existence and the uniqueness of the optimal trading strategy. It also ensures the smoothness of the value function except for a set of measure zero.

**Theorem 4.3.1.** *(i) The HJB equation* (4.29)–(4.32) *admits a unique viscosity solution, and the value function is the viscosity solution.*

*(ii) The value function is* $C^{2,2,1}$ *in* $(x, y) \in \mathcal{S}_d \setminus (\{y = 0\} \cup \{x = 0\})$, $t \in (t_{2i}, t_{2i+1})$ *and in* $(x, y) \in \mathcal{S}_n$, $t \in (t_{2i-1}, t_{2i})$, *for* $i = 0, 1, ..., N$.

*Proof.* Part (i) can be proved using a similar argument as in Shreve and Soner (1994). To show part (ii), we can follow Dai and Yi (2009) to reduce the HJB equation to a double obstacle problem in the day time $(t_{2i}, t_{2i+1})$. Then we can obtain $C^{2,2,1}$ smoothness of the value function for $t \in (t_{2i}, t_{2i+1})$. The smoothness function of the value function in the night time $(t_{2i-1}, t_{2i})$ is apparent. $\square$

The homogeneity of the utility function $U(\cdot)$ and the fact that $\mathcal{A}(\rho x, \rho y) = \rho\mathcal{A}(x, y)$ for all $\rho > 0$ imply that $V + \frac{1}{1-\gamma}$ is concave in $(x, y)$ and homogeneous of degree $1 - \gamma$ in $(x, y)$ [cf. Fleming and Soner (1993), Lemma VIII.3.2]. This homogeneity implies that

$$V(x, y, t) = y^{1-\gamma}\phi\left(\frac{x}{y}, t\right) - \frac{1}{1-\gamma}, \tag{4.33}$$

for some function $\phi : (\alpha - 1, \infty) \times [0, T] \to \mathbb{R}$. As in the case of previous chapter, due to the fact that the risk premium is positive, short sale is never optimal and thus we can confine the problem to the domain where $y > 0$. Let

$$z = \frac{x}{y} \tag{4.34}$$

denote the ratio of the dollar amount invested in the bank account to that in the stock.

By the same arguments as in Dai and Yi (2009), we are able to show that in the "daytime" when trading is allowed, the solvency region $\Omega_z \equiv (\alpha - 1, \infty) \times [0, T)$ splits

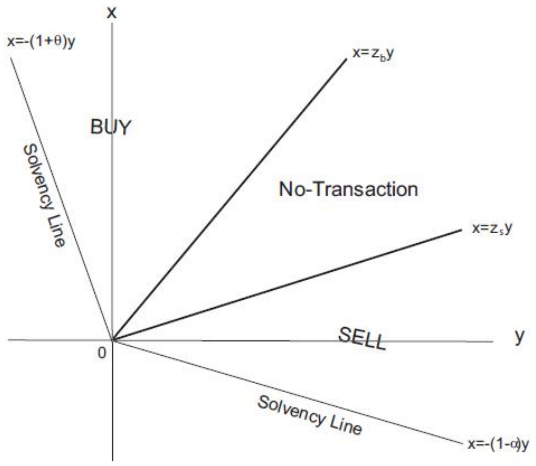



Figure 4.1: The Solvency Region

to three parts, the selling region (SR), the buying region (BR), and the no transaction region (NT). They can be characterized by two free boundaries $z_s^*(t)$ and $z_b^*(t)$, such that

$$\begin{aligned} SR &= \{(z,t) \in \Omega_z, \ \ z \le z_s^*(t)\}, \\ BR &= \{(z,t) \in \Omega_z, \ \ z \ge z_b^*(t)\}, \\ NT &= \{(z,t) \in \Omega_z, \ \ z_s^*(t) < z < z_b^*(t)\}. \end{aligned}$$

A time snapshot of these regions is depicted in Figure 4.1.

The optimal trading strategy during "daytime" $(t_{2i}, t_{2i+1})$ is to transact a minimum amount of the stock to keep the ratio $z_t$ in the optimal no-transaction region. Therefore the determination of the optimal trading strategy reduces to the determination of the optimal no-transaction region, or equivalently, the two trading boundaries: $z_b^*(t)$ and $z_s^*(t)$. In contrast to the no-transaction-costs case, the optimal fractions of the liquidated wealth invested in both the bond and the stock change stochastically, since $z_t$ varies stochastically due to no transaction in NTR.

So far the arguments for the "daytime" evolution are almost the same as in Dai and Yi (2009). The difference brought forward by market closure mechanism is revealed in next proposition, where we need a connection condition at market close $t_{2i+1}$. This condition is implied by (4.31).

**Proposition 9.** *There exist* $z_s^*(t_{2i+1}) \in [-(1-\alpha), \infty)$ *and* $z_b^*(t_{2i+1}) \in (-(1-\alpha), \infty]$ *such that* $V(x, y, t_{2i+1})$ *is given as follows:*

$$\begin{aligned} &V(x, y, t_{2i+1}) = V\left(x, y, t_{2i+1}^+\right) && z_s^*(t_{2i+1}) < x/y < z_b^*(t_{2i+1}) \\ &-(1-\alpha)V_x(x, y, t_{2i+1}) + V_y(x, y, t_{2i+1}) = 0 && x/y \le z_s^*(t_{2i+1}) \\ &(1+\theta)V_x(x, y, t_{2i+1}) - V_y(x, y, t_{2i+1}) = 0 && x/y \ge z_b^*(t_{2i+1}). \end{aligned} \tag{4.35}$$

*Proof.* By definition, the value function $V(x, y, t)$ is concave in $x$ and $y$. Then we can deduce that the following two domains

$$\begin{aligned} E_b &\equiv \left\{(x, y) : (1+\theta) V_x - V_y|_{t=t_{2i1}^+} > 0,\ x > 0,\ y > 0\right\} \\ E_s &\equiv \left\{(x, y) : -(1-\alpha) V_x + V_y|_{t=t_{2i1}^+} > 0,\ x > 0,\ y > 0\right\} \end{aligned}$$

must be connected. Here we confine to $x > 0$ and $y > 0$, in order to ensure solvency. Due to the homogeneity of the value function, we can define $z_b^*(t_{2i+1})$ and $z_s^*(t_{2i+1})$ as

$$z_b^*(t_{2i+1}) \equiv \sup\left\{\frac{x}{y} : (x, y) \in E_b\right\}, \tag{4.36}$$

$$z_b^*(t_{2i+1}) \equiv \sup\left\{\frac{x}{y} : (x, y) \in E_s\right\}. \tag{4.37}$$

Then we consider (4.31) and look for the first order condition for the maximization. For any $\Delta > 0$, we have

$$\begin{aligned} \frac{d}{d\Delta}V\left(x - (1+\theta)\Delta, y + \Delta, t_{2i+1}^+\right) &= -(1+\theta) V_x + V_y, \\ \frac{d}{d\Delta}V\left(x + (1-\alpha)\Delta, y - \Delta, t_{2i+1}^+\right) &= (1-\alpha) V_x - V_y. \end{aligned}$$

combining with (4.36) and (4.37), we get the desired result. □

To investigate in the $(z, t)$ plane, it is straightforward to verify that by transformation

(4.33), equations (4.29), (4.30) and (4.32) reduce to the following system:

$$\begin{cases} \max\left\{\phi_t+\mathcal{L}_1\phi,(z+1-\alpha)\phi_z-(1-\gamma)\phi,-(z+1+\theta)\phi_z+(1-\gamma)\phi\right\}=0, & t\in[t_{2i},t_{2i+1}) \\ \phi_t+\mathcal{L}_1\phi=0, & t\in(t_{2i-1},t_{2i}) \\ \phi(z,t_{2i+1})=\max_{k\in\hat{C}(z)}(1+k)^{1-\gamma}\phi\left(\frac{z-(1+\theta)k^{+}+(1-\alpha)k^{-}}{1+k},t_{2i+1}^{+}\right) \\ \phi(z,T)=\frac{1}{1-\gamma}(z+1-\alpha)^{1-\gamma}, \end{cases}$$

where

$$\mathcal{L}_1\phi=\frac{1}{2}\sigma(t)^2z^2\phi_{zz}+\beta_2(t)z\phi_z+\beta_1(t)\phi,$$
$$\hat{C}(z)=\{k\geq-1:z-(1+\theta)k^{+}+(1-\alpha)k^{-}\geq0\},$$

with $\beta_1(t)=(1-\gamma)\left(\mu(t)-\frac{1}{2}\gamma\sigma(t)^2\right)$ and $\beta_2(t)=-\left(\mu(t)-r-\gamma\sigma(t)^2\right)$. The solvency region $\mathcal{S}_d$ in the original $(x,y,t)$ space becomes $(-(1-\alpha),\infty)\times[0,T)\equiv\mathcal{S}_d^z$ in the $(z,t)$ plane, while $\mathcal{S}_n$ becomes $[0,\infty)\times[0,T)\equiv\mathcal{S}_n^z$. The connection conditions (4.35) at $t_{2i+1}$ turns into

$$\begin{cases} \phi\left(z,t_{2i+1}\right)=\phi\left(z,t_{2i+1}^{+}\right), & z_s^*\left(t_{2i+1}\right)<z<z_b^*\left(t_{2i+1}\right), \\ -(z+1-\alpha)\phi_z\left(z,t_{2i+1}\right)+(1-\gamma)\phi\left(z,t_{2i+1}\right)=0, & z\leq z_s^*\left(t_{2i+1}\right), \\ (z+1+\theta)\phi_z\left(z,t_{2i+1}\right)-(1-\gamma)\phi\left(z,t_{2i+1}\right)=0, & z\geq z_b^*\left(t_{2i+1}\right). \end{cases}$$

### 4.3.2 Behaviors of the free boundaries

The nonlinearity of the HJB equation and the time-varying nature of the free boundaries make it difficult to investigate behaviors of free boundaries directly. Instead, as in previous chapter or as in Dai and Yi (2009), we would like to transform the above problem into a double obstacle problem, which is much easier to analyze.

In the sense of bank account-to-stock ratio, we let

$$z_M=\frac{\gamma\sigma_d^2}{\mu_d-r}-1$$

be the daytime Merton line. The following comparative statics for trading boundaries should hold.

**Proposition 10.** *For any $t \in [t_{2i}, t_{2i+1})$, we have*

*(i)* $z_b^*(t) \geq (1+\theta) z_M$;

*(ii)* $z_s^*(t) \leq (1-\alpha) z_M$.

*Proof.* The proof is actually based on the approach used in previous chapter. And the absence of consumption makes it much easier to deal with.

First we take the following transformations:

$$\begin{aligned} w(z,\tau) &= \frac{1}{1-\gamma} \log\left((1-\gamma)\,\phi(z,t)\right) \\ \tau &= T - t \end{aligned}$$

then we get a double obstacle problem:

$$\begin{cases} \min\left\{-w_t - \mathcal{L}_2 w,\ \frac{1}{z+1-\alpha} - w_z = 0,\ w_z - \frac{1}{z+1+\theta}\right\}, & t \in [t_{2i+1}, t_{2i+2}) \\ -w_t - \mathcal{L}_2 w = 0, & t \in (t_{2i}, t_{2i+1}) \\ v(z,T) = \log(z+1-\alpha) & \end{cases} \tag{4.38}$$

with the connection condition

$$\begin{cases} w(z, t_{2i+1}) = w\left(z, t_{2i}^+\right) & z_s^*(t_{2i+1}) < z < z_b^*(t_{2i+1}) \\ w_z(z, t_{2i+1}) = \frac{1}{z+1-\alpha} & z \leq z_s^*(t_{2i+1}) \\ w_z(z, t_{2i+1}) = \frac{1}{z+1+\theta} & z \geq z_b^*(t_{2i+1}). \end{cases} \tag{4.39}$$

To remove the constraints on gradients in (4.38), we take one more step to let

$$v(z,t) = w_z(z,t)$$

then following Dai and Yi (2009), we are able to show that $v$ satisfies the following parabolic double obstacle problem:

$$\begin{cases} \max\left\{\min\left\{-v_t - \mathcal{L} v,\ v - \frac{1}{z+1+\theta}\right\},\ \frac{1}{z+1-\alpha} - v\right\} = 0, & t \in [t_{2i+1}, t_{2i+2}) \\ -v_t - \mathcal{L} v = 0, & t \in (t_{2i}, t_{2i+1}) \\ v(z,T) = \frac{1}{z+1-\alpha} & \end{cases} \tag{4.40}$$

subject to the connection condition

$$\begin{cases} v\left(z,t_{2i+1}\right)=v\left(z,t_{2i}^{+}\right) & z_s^*\left(t_{2i+1}\right)<z<z_b^*\left(t_{2i+1}\right) \\ v\left(z,t_{2i+1}\right)=\frac{1}{z+1-\alpha} & z\le z_s^*\left(t_{2i+1}\right) \\ v\left(z,t_{2i+1}\right)=\frac{1}{z+1+\theta} & z\ge z_b^*\left(t_{2i+1}\right). \end{cases}$$

we then infer that for any $t\in\left(t_{2i+1},t_{2i+2}\right)$,

$$\begin{aligned} (\mathbf{SR})_t &\equiv \left\{z:v\left(z,t\right)=\frac{1}{z+1-\alpha}\right\}=\left\{z\le z_s^*\left(t\right)\right\} \\ (\mathbf{BR})_t &\equiv \left\{z:v\left(z,t\right)=\frac{1}{z+1+\theta}\right\}=\left\{z\ge z_b^*\left(t\right)\right\} \end{aligned}$$

Thanks to (4.40), we have

$$\begin{aligned} \left(-\frac{\partial}{\partial t}-\mathcal{L}\right)\left(\frac{1}{z+1-\alpha}\right) &\le 0 \quad \text{for} \quad z\in(\mathbf{SR})_t \ \text{(i.e. } z\le z_s^*\left(t\right)), \\ \left(-\frac{\partial}{\partial t}-\mathcal{L}\right)\left(\frac{1}{z+1+\theta}\right) &\ge 0 \quad \text{for} \quad z\in(\mathbf{BR})_t \ \text{(i.e. } z\ge z_b^*\left(t\right)). \end{aligned}$$

Note that

$$\begin{aligned} \left(-\frac{\partial}{\partial t}-\mathcal{L}\right)\left(\frac{1}{z+1-\alpha}\right) &= -\mathcal{L}\left(\frac{1}{z+1-\alpha}\right) \\ &= \frac{\left(1-\alpha\right)\left(\mu_d-r\right)}{\left(z+1-\alpha\right)^3}\left[z+\left(1-\alpha\right)\frac{\mu_d-r-\gamma\sigma_d^2}{\mu_d-r}\right] \\ &= \frac{\left(1-\alpha\right)\left(\mu_d-r\right)}{\left(z+1-\alpha\right)^3}\left[z-\left(1-\alpha\right)z_M\right] \\ &\le 0 \end{aligned}$$

implies

$$z\le\left(1-\alpha\right)z_M,\ \forall z\in(\mathbf{SR})_t\,,$$

thus

$$z_s^*\left(t\right)\le\left(1-\alpha\right)z_M,$$

and similarly

$$\left(-\frac{\partial}{\partial t}-\mathcal{L}\right)\left(\frac{1}{z+1+\theta}\right)=\frac{\left(1+\theta\right)\left(\mu_d-r\right)}{\left(z+1+\theta\right)^3}\left[z-\left(1+\theta\right)z_M\right]$$

leads to the conclusion that

$$z_b^*\left(t\right)\ge\left(1+\theta\right)z_M.$$

The proof is completed. □

Furthermore, to investigate how market closure impacts the trading boundaries, we are interested in the behavior of trading boundaries right at the time instant when market closes, i.e., from $t = t_{2i+1}^{-}$ to $t = t_{2i+1}$.

**Proposition 11.** *At the time instances just before market closes, it holds that*

$$z_s^*\left(t_{2i+1}^-\right) = \min\left\{z_s^*(t_{2i+1}), (1-\alpha)\, z_M\right\}; \tag{4.41}$$

$$z_b^*\left(t_{2i+1}^-\right) = \max\left\{z_b^*(t_{2i+1}), (1+\theta)\, z_M\right\}, \tag{4.42}$$

*Proof.* We would like to prove only (4.41). The proof of (4.42) is straightforward following the same idea and steps.

First, let us show

$$z_s^*\left(t_{2i+1}^-\right) \le z_s^*\left(t_{2i+1}\right).$$

To make use of contradiction, we first suppose not, i.e., $z_s^*\left(t_{2i+1}^-\right) > z_s^*\left(t_{2i+1}\right)$.

Let $w(z,t)$ be the solution to the problem (4.38). Since $\left(z_s^*\left(t_{2i+1}^-\right), t_{2i+1}\right)$ is in the no-transaction region, $w(z,t)$ is continuous at $\left(z_s^*\left(t_{2i+1}^-\right), t_{2i+1}\right)$, namely,

$$w\left(z_s^*\left(t_{2i+1}^-\right), t_{2i+1}^-\right) = w\left(z_s^*\left(t_{2i+1}^-\right), t_{2i+1}\right).$$

Then for $z \in \left(z_s^*\left(t_{2i+1}\right),\ z_s^*\left(t_{2i+1}^-\right)\right)$, it holds that

$$\begin{aligned} w\left(z,\, t_{2i+1}^-\right) &= w\left(z_s^*\left(t_{2i+1}^-\right), t_{2i+1}^-\right) - \int_z^{t_{2i+1}^-} \frac{1}{\xi + 1 - \alpha} d\xi \\ &< w\left(z_s^*\left(t_{2i+1}^-\right), t_{2i+1}^-\right) - \int_z^{t_{2i+1}^-} w\left(\xi, t_{2i+1}\right) d\xi \\ &= w\left(z,\, t_{2i+1}\right) \end{aligned}$$

which contradicts to the connection condition (4.39).

Second, it is clear that $z_s^*\left(t_{2i+1}^-\right) \le (1-\alpha)\, z_M$ from Proposition 10. So now we deduce that

$$z_s^*\left(t_{2i+1}^-\right) \le \min\left\{z_s^*\left(t_{2i+1}\right),\ (1-\alpha)\, z_M\right\}.$$

Furthermore, if $z_s^*\left(t_{2i+1}^-\right) < \min\left\{z_s^*\left(t_{2i+1}\right),\ (1-\alpha)\, z_M\right\}$, then for

$$z \in \left(z_s^*\left(t_{2i+1}^-\right),\ \min\left\{z_s^*\left(t_{2i+1}\right),\ (1-\alpha)\, z_M\right\}\right),$$

we have

$$v\left(z, t_{2i+1}\right) = \frac{1}{z+1-\alpha},$$

and the equation below is satisfied:

$$-v_t - \mathcal{L}v|_{(z,t_{2i+1})} = 0$$

It follows that

$$\begin{aligned} v_t|_{(z,t_{2i+1})} &= -\mathcal{L}\left(\frac{1}{z+1-\alpha}\right) \\ &= \frac{(1-\alpha)\left(\mu_d - r\right)}{(z+1-\alpha)^3}\left[z-(1-\alpha)\, z_M\right] \\ &< 0, \end{aligned}$$

which conflicts with the fact that

$$v_t|_{(z,t_{2i+1})} \geq 0.$$

So the contradiction assumption must not be true. The proof is completed. □

When market closes, an investor should adjust his portfolio to be within the interval $[z_s^*(t_{2i+1}), z_b^*(t_{2i+1})]$. Proposition 11 suggests that an investor may optimally wait until the market closing time to adjust his portfolio. For example, in the case $z_s^*\left(t_{2i+1}^-\right) = (1-\alpha)z_M < z_s^*(t_{2i+1})$, if the investor's position is on the sell boundary $z_s^*\left(t_{2i+1}^-\right)$ right before market closes, he will perform a discrete sale to adjust his portfolio to $z_s^*(t_{2i+1})$. Similarly, an investor may make a discrete purchase to adjust his portfolio to $z_b(t_{2i+1})$. This is consistent with the empirical evidence that trading volume increases at market close.

## 4.4 Analysis

In this section we provide some numerical analysis on the impact of market closure and time-varying return dynamics on optimal trading strategy, the liquidity premia, the loss from market closure, and the loss from adopting the "optimal" strategy implied by the standard assumption of continuously open market and constant return dynamics.

### 4.4.1 Liquidity premium

Most of the existing literature found that transaction costs had a second order effect on risk premium. For example, the seminal work of Constantinides (1986) showed that for a 1% proportional transaction cost rate, an investor only needed about 0.1% compensation in risk premium (i.e., the liquidity premium is only about 0.1%). The main intuition behind this result is that investor does not need to trade much given that the investment opportunity set (such as expected return and volatility) as assumed in Constantinides (1986) is constant. Thus the investor pays not much transaction cost and it naturally leads to low liquidity premium. If the investor's opportunity set varies with time, one may infer that an investor needs to trade more and thus requires higher compensation for transaction cost. Indeed, Jang et. al. (2007) showed that when there were two regimes with different volatilities, then the transaction cost could have a much higher effect on liquidity premium. However, due to the infrequency of regime switching and the small difference in volatilities across regimes, the effect was still small. For example, the liquidity premium to transaction cost ratio (LPTC) only increased from 0.1 to about 0.5 in most cases in Jang et. al. (2007). With periodic market closure, the investor's opportunity set would change much more frequently. In this subsection, we show that incorporating market closure and the significant difference of volatilities across trading and non-trading period can make transaction cost have a first order effect on liquidity premium. In other words, the liquidity premium to transaction cost ratio can be well above 1.

We begin the numerical illustration by selecting the benchmark market first. Consider Market M, which is continuously open. Assume there is no transaction cost and no time-varying stock return dynamics in Market M (the Merton's case), then we can take the value function in Market M as benchmark. Let Market A be the actual market with positive transaction costs, different volatilities across trading and non-trading period, and periodic market closure. Given that the expected return $\mu_d = \mu_n = \mu$, we denote the time 0 value functions in Market M and Market A respectively by $V_M\left(x, y, 0; \mu\right)$ and

$V_A(x, y, 0; \mu, \alpha)$. Following Constantinides (1986), we solve the equation

$$V_M(z_M, 1, 0; \mu - \delta) = V_A(z_M, 1, 0; \mu, \alpha) \tag{4.43}$$

for the liquidity premium $\delta$ that measures how much an investor is willing to give up in risk premium to avoid transaction cost, when he starts at the daytime Merton line $z_M$. The $\delta$ from equation (4.43) is affected by the time varying volatility and the inability to trade overnight in Market A. To separate the two effects, we also take another market Market B, which is exactly the same as Market A except that there is no market closure so investor can trade overnight subject to the same daytime transaction costs. Denote the value function in Market B by $V_B(x, y, 0; \mu, \alpha)$, the following equation (4.44) will solve liquidity premium $\tilde{\delta}$ in another sense, which measures the compensation for transaction costs in our model due to time varying return dynamics only.

$$V_M(z_M, 1, 0; \mu - \delta) = V_B(z_M, 1, 0; \mu, \alpha) \tag{4.44}$$

In general, the effect of transaction cost on liquidity premium comes from two sources. One is the direct transaction cost payment incurred at each trade. The other source is the adoption of suboptimal trading strategy. Here by "suboptimal" we mean that although the trading strategy is optimal when transaction costs apply, it is suboptimal in the absence of transaction costs. To understand which source is the main driving force for the liquidity premium, we also compute the liquidity premium caused by the suboptimal trading strategy alone. Specifically, let $(I, D)$ be the optimal trading strategy in the presence of transaction costs in Market A, and $V_M^{(I,D)}(x, y, 0; \mu)$ be the value function from following the strategy $(I, D)$ in Market M (where actually no transaction costs apply). Then we solve

$$V_M(z_M, 1, 0; \mu - \delta^0) = V_M^{(I,D)}(z_M, 1, 0; \mu) \tag{4.45}$$

for the liquidity premium $\delta^0$, which is due to the suboptimality of the trading strategy $(I, D)$ in Market M.

For simplicity, we assume from now on that every day market opens for $\Delta_d = 6.5$ hours (from 9:30am to 4pm) and closes for $\Delta_n = 24 - 6.5 = 17.5$ hours. Let the average volatility be $\sigma$ and the ratio of the day volatility to night volatility be $k \equiv \sigma_d/\sigma_n$. Solving the equations

$$\begin{cases} \sigma_d^2 \Delta_d + \sigma_n^2 \Delta_n = \sigma^2 \left( \Delta_d + \Delta_n \right), \\ \sigma_d = k \sigma_n, \end{cases}$$

gives

$$\begin{cases} \sigma_d = k\sigma \cdot \sqrt{\frac{\Delta t_d + \Delta t_n}{k^2 \Delta t_d + \Delta t_n}}, \\ \sigma_n = \sigma \cdot \sqrt{\frac{\Delta t_d + \Delta t_n}{k^2 \Delta t_d + \Delta t_n}}. \end{cases} \tag{4.46}$$

To make the closest possible comparison with Constantinides (1986), we set the default parameter values at $\mu_d = \mu_n = \mu = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\alpha = 1\%$, $\theta = 1\%$, $k = 3$, and $T = 10$. Although both $\mu$ and $r$ may be high relative to realizations in recent years, what matters for our analysis is the risk premium $(\mu - r)$. Besides, the existing literature on intraday price dynamics found that an average per-hour ratio of day-time to overnight volatility was around 4.0 and that the expected returns were not significantly different across day and night (e.g., Stoll and Whaley (1990), Lockwood and Linn (1990), Tsiakas (2008)). Our choice of a smaller default value $k = 3$ biases against us in finding significant effects of market closure.

In Table 4.1 we compare the optimal no-transaction boundaries and the LPTC ratios in our model with those reported by Constantinides (1986). This table shows that the LPTC ratios are much higher. In fact, for a reasonable transaction cost of $< 1\%$ each way for trading stock index such as S&P 500, transaction costs can have more than a first order effect. For example, at $\alpha = \theta = 0.5\%$, the LPTC ratio is as high as 3.53, more than 20 times higher than what is found by Constantinides (1986). This magnitude of LPTC ratio is consistent with empirical findings such as those by Amihud and Mendelson (1986) with a LPTC ratio of 2.4.

Table 4.2 compares the liquidity premium from different sources with the total liquidity premium. The first panel shows that when the investor can trade overnight with the same transaction cost rates as incurred in daytime. It suggests that the effect of the inability to trade overnight on liquidity premium is negligible. The difference of volatilities across day and night is overwhelmingly dominating the high liquidity premium.

Table 4.1: Optimal policy and liquidity premia against transaction cost rates

| $\alpha = \theta =$: | 0.005 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.10 | 0.15 |
|---|---|---|---|---|---|---|---|---|
| $z_b^*\left(t_1^-\right)$ | 3.567 | 3.585 | 3.621 | 3.656 | 3.692 | 3.727 | 3.905 | 4.089 |
| $z_s^*\left(t_1^-\right)$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $z_b^*\left(t_1\right)$ | 0.759 | 0.813 | 0.909 | 1.009 | 1.120 | 1.242 | 2.132 | 4.061 |
| $z_s^*\left(t_1\right)$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| $\delta/\alpha$ | 3.68 | 1.90 | 1.01 | 0.71 | 0.56 | 0.47 | 0.28 | 0.21 |
| $\delta/\delta_C$ | 23.01 | 13.56 | 8.05 | 5.73 | 4.54 | 3.82 | 2.14 | 1.44 |
| *Constantinides* (1986) | | | | | | | | |
| $z_{b,\,C}^*$ | 0.690 | 0.726 | 0.783 | 0.832 | 0.877 | 0.920 | 1.122 | 1.326 |
| $z_{s,\,C}^*$ | 0.566 | 0.561 | 0.555 | 0.550 | 0.546 | 0.542 | 0.525 | 0.509 |
| $\delta_C/\alpha$ | 0.16 | 0.14 | 0.13 | 0.12 | 0.12 | 0.12 | 0.13 | 0.14 |

$z_b^*$ and $z_s^*$ are the buying and selling boundaries. $t_1^-$ is just before first closing and $t_1$ is at first closing. $\delta$ and $\delta_C$ are the time 0 liquidity premiums starting from the daytime Merton line. Other parameters: $T = 10$, $\mu_d = \mu_n = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\Delta_d = 6.5$ hours, $\Delta_n = 17.5$ hours, $k = 3$, and $\gamma = 2$.

Therefore market closure per se is not important for our results, it is the large volatility variation caused by market closure that significantly raises the liquidity premium. This finding is consistent with Jang et. al. (2007). In their paper, they also found that the higher liquidity premium (compared to Constantinides (1986)) came from volatility difference across the bear regime and the bull regime. However, since the frequency of regime switching is low and the empirically found volatility difference across the two regimes is small, the typical LPTC ratio in Jang et. al. (2007) was around 0.5, which was still insufficient to account for the empirical evidence.

One typical explanation for a higher liquidity premium when investment opportunity set changes is the increase in trading frequency and transaction cost payment. To help understand whether higher transaction cost payment is the main driving force behind

Table 4.2: Sources of higher liquidity premium

| $\alpha = \theta =$: | 0.005 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.10 | 0.15 |
|---|---|---|---|---|---|---|---|---|
| | | | *This Model without Market Closure* | | | | | |
| $\tilde{\delta}/\alpha$ | 3.67 | 1.90 | 1.00 | 0.71 | 0.56 | 0.46 | 0.28 | 0.21 |
| | | | *This Model with Market Closure* | | | | | |
| $\delta/\alpha$ | 3.68 | 1.90 | 1.01 | 0.71 | 0.56 | 0.47 | 0.28 | 0.21 |
| $\delta^0/\delta$ | 98% | 95% | 90% | 86% | 83% | 81% | 76% | 82% |
| | | | *Constantinides* (1986) | | | | | |
| $\delta_C^0/\delta_C$ | 10% | 14% | 20% | 24% | 27% | 30% | 36% | 36% |

the high LPTC ratio in our model, we report the liquidity premium $\delta^0$ due to the suboptimality of trading strategy in the second panel of Table 4.2. In contrast to conventional wisdom, it turns out that only a small percentage of the liquidity premium is from direct transaction cost payment. The vast majority of the liquidity premium comes from the suboptimal portfolio position compared to no transaction cost case. This finding suggests that with the large volatility difference across trading and non-trading period, the investor choose to widen up his no transaction region to reduce his transaction cost payment.Indeed, as Table 4.1 shows, the no-transaction region in this model is much wider than that in Constantinides (1986). For example, when $\alpha = \theta = 0.01$, the time 0 no transaction region in the market closure model is $(0.430, 3.608)$ which is significantly wider than $(0.561, 0.726)$ that is optimal in Constantinides (1986).

One thing to note is that wider no transaction regions in our model do not necessarily lead to lower trading frequencies than that in Constantinides (1986). Since frequent market closure may increase rebalancing needs and thus trading frequency as well. To compare the trading frequency and transaction cost payment across our model and Constantinides (1986)' model, we conduct Monte Carlo simulations in these two cases and report related results in Table 4.3.

Table 4.3: Simulated trading frequency and transaction costs

| | This model | | *Constantinides* (1986) | |
|---|---|---|---|---|
| $\alpha = \theta =$ | 0.005 | 0.01 | 0.005 | 0.01 |
| Sell/Buy ($ ratio) | 2.8293 | 3.3010 | 41.9627 | 35.8950 |
| Sell/Buy (share ratio) | 0.9127 | 0.9122 | 15.3734 | 12.1396 |
| Average BS time | 2.8077 | 4.4521 | 1.8375 | 2.6203 |
| # of purchases p.a. | 7.55 | 5.85 | 17.72 | 19.21 |
| # of sales p.a. | 15.93 | 13.80 | 596.71 | 540.47 |
| Trading frequency | 23.48 | 19.65 | 614.42 | 559.68 |
| PVTC ($) | 0.52 | 0.91 | 0.23 | 0.41 |

PVTC is the discounted transaction costs paid as a percentage of the initial wealth. Other parameters: $T = 10$, $\mu_d = \mu_n = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\Delta_d = 6.5$ hours, $\Delta_n = 17.5$ hours, $k = 3$, and $\gamma = 2$.

Table 4.3 shows that the trading frequency in Constantinides (1986) is much higher (almost 30 times) than that in the market closure model. This confirms the intuition that to avoid large transaction cost payment, the investor chooses a trading strategy to significantly reduce trading frequency. On the other hand, Table 4.3 also shows that even though the trading frequency is much lower, the transaction costs paid in this model is still higher than that in Constantinides (1986). For example, with 1% transaction cost rate, the present value of transaction costs paid is 0.91% of the initial wealth while it is only 0.41% in Constantinides (1986). This is mainly because trading in this model can involve large discrete trading at market close and market open, while in Constantinides (1986), only infinitesimal trading at the boundaries can happen. In other words, the average per-trade trading size is much larger in this model, which is also corroborated by the trading numbers reported in Table 4.3. It is also suggested that the investor sells more often than buys. This is simply because stock price goes up on average.

In Figure 4.2, we plot the LPTC ratios against the day-night volatility ratio $k$ for

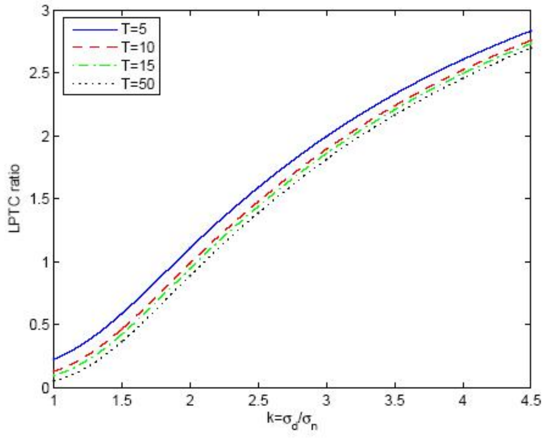



Figure 4.2: LPTC ratios against day-night volatility ratio $k$.

Parameter values: $\mu_d = \mu_n = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\gamma = 2$, $\Delta_d = 6.5$ hours, $\Delta_n = 17.5$ hours, $\alpha = \theta = 0.01$.

three different investment horizons of $T = 5$, 10, and 15 years. This figure shows that LPTC is sensitive to and increasing in the difference between daytime and overnight volatility. For example, at $k = 2$, the LPTC ratio is about 0.99 and it increases to 1.83 when $k$ increases to 3. It is worth noting that at $k = 1$, the LPTC ratio is close to that of Constantinides (1986). This suggests that the main reason for the large impact of transaction costs on liquidity premium is the time-varying stock return volatility, not the market closure in itself.

Figure 4.2 also illustrates how the liquidity premium behaves as the investment horizon changes. On the one hand, the LPTC ratio increases as the investment horizon decreases. This is because that the investor needs to liquidate stock positions sooner with a shorter horizon. On the other hand, the high LPTC ratio in this model is not mainly due to the finite investment horizon. We can see from the figure that even with a long horizon of $T = 50$, the LPTC ratios for large $k$ are still well above 1.

Table 4.4 reports optimal no-transaction boundaries and liquidity premia against risk aversion coefficient $\gamma$ for two different transaction cost rates $\alpha = \theta = 1\%$. This table shows that the LPTC ratio is more than 10 times higher than that in Constantinides (1986) and that transaction cost has more than a first order effect for a range of reasonable risk aversion levels. In addition, LPTC ratio increases with risk aversion. Intuitively, as risk aversion increases, an investor invests less in the stock and therefore he is willing to give up more risk premium in exchange for 0 transaction cost.

Table 4.4: Optimal policy and liquidity premia against risk aversion coefficients

| | $\gamma$ | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 |
| $z_b^*(0)$ | 3.608 | 5.922 | 8.248 | 10.546 | 12.864 |
| $z_s^*(0)$ | 0.430 | 1.067 | 1.712 | 2.360 | 3.009 |
| $z_b^*(t_1^-)$ | 3.585 | 5.889 | 8.180 | 10.480 | 12.788 |
| $z_s^*(t_1^-)$ | 0.000 | 0.000 | 0.038 | 0.295 | 0.552 |
| $z_b^*(t_1)$ | 0.813 | 1.846 | 2.878 | 3.910 | 4.942 |
| $z_s^*(t_1)$ | 0.000 | 0.000 | 0.011 | 0.261 | 0.511 |
| $\delta/\alpha$ | 1.832 | 1.910 | 1.918 | 1.923 | 1.926 |
| $\delta/\delta_C$ | 13.087 | 11.936 | 11.282 | 10.683 | 10.139 |
| | *Constantinides* (1986) | | | | |
| $z_b^*(0)$ | 0.726 | 1.736 | 2.747 | 3.759 | 4.785 |
| $z_s^*(0)$ | 0.561 | 1.304 | 2.045 | 2.778 | 3.521 |
| $\delta_C/\alpha$ | 0.14 | 0.16 | 0.17 | 0.18 | 0.19 |

$z_b^*$ and $z_s^*$ are the buying and selling boundaries. $t_1^-$ is just before first closing and $t_1$ is at first closing. $\delta$ and $\delta_C$ are the time 0 liquidity premiums starting from the daytime Merton line. Other parameters: $T = 10$, $\mu_d = \mu_n = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\Delta_d = 6.5$ hours, $\Delta_n = 17.5$ hours, $k = 3$, and $\alpha = \theta = 0.01$.

### 4.4.2 The loss from ignoring volatility variation

There exists an extensive literature on the intraday volatility and expected return dynamics. One of the most robust results is that the stock return volatility is much higher when market is open than when it is closed, while the expected returns are not significantly different across the two periods. However, most of the standard literature (e.g., Merton (1987)) assumes that market is continuously open with a constant volatility. In this subsection we show that using the "optimal" trading strategy derived under this assumption implies large wealth loss for the investor.

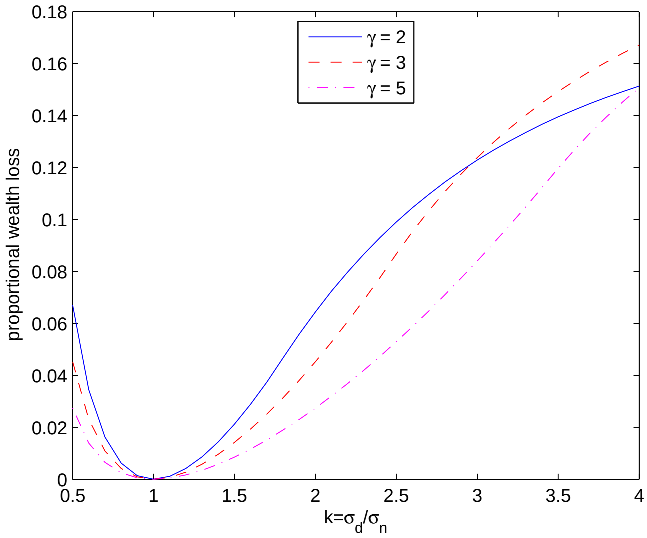
Consider Market C where trading is continuously allowed and the stock has a constant return volatility of $\sigma$ across trading and non-trading period. Let $\pi_t^C$ be the optimal trading strategy in Market C. Suppose the actual market is Market A, where market closes at night and the stock has a daytime return volatility of $\sigma_d$ and a night time return volatility of $\sigma_n$. We examine the cost for the investor from following strategy $\pi_t^C$ in Market A in terms of wealth loss. Specifically, let $V_A(x, y, 0)$ be the value function in Market A following the correct strategy regarding varying stock return volatility, and $V_A^{\pi_C}(x, y, 0)$ be the value functions in Market A following the wrong strategy $\pi_t^C$, respectively. Then we solve

$$V_A(1 - \Delta, 0, 0) = V_A^{\pi_C}(1, 0, 0)$$

for $\Delta$ that measures the percentage of initial wealth an investor is willing to give up in order to use the correct strategy. The explicit expressions for the value functions in the no-transaction-cost case are provided in Section 4.2.3.

In Figure 4.3 we plot the wealth loss from following the wrong strategy against $k$ in the absence of transaction costs for three different levels of risk aversion: $\gamma = 2,\ 3,\ 5$. This figure shows that following the optimal strategy proposed by the standard models is costly. For example, at $k = 3$, for an investor with a risk aversion coefficient of 2, the loss is as high as 12.29% of his initial wealth. Figure 4.3 also shows that the wealth loss increases as the day-night volatility difference increases, which is the natural implication

Figure 4.3: Wealth loss from following standard strategy against day-night volatility ratio $k$.




Parameter default values: $T = 10$, $\mu_d = \mu_n = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\Delta_d = 6.5$ hours, $\Delta_n = 17.5$ hours, $\alpha = \theta = 0.01$.

of the assumption of constant volatility in Market C. Interestingly, while the wealth loss for a more risk averse investor is lower when the day-night volatility ratio is low, it may be higher if the ratio is high. For example, at $k = 3$, the wealth loss for an investor with a risk aversion coefficient of 3, the loss is 12.38% of his initial wealth. Intuitively, an investor overinvests (underinvests) during market open if and only if the day-night volatility ratio $k > 1$ ($k < 1$). A more risk averse investor overinvests less during market open and underinvests more during market close. If the day-night volatility ratio is high, the more severe underinvestment during night dominates the reduction of overinvestment during day and thus a more risk averse investor incurs a greater loss.

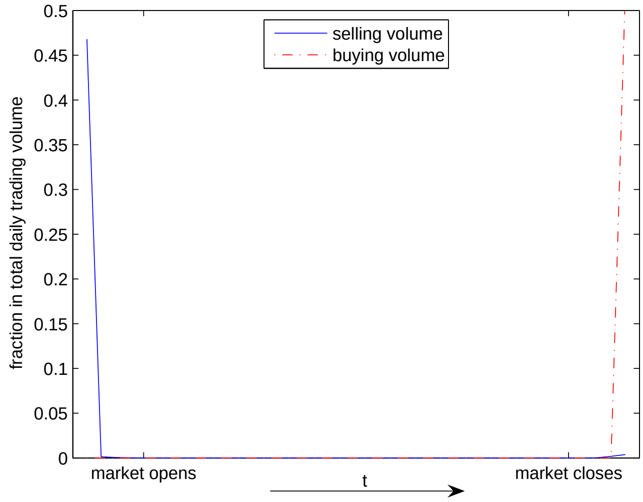



Figure 4.4: The distribution of the fraction of total trading volume across trading time. Parameter values: $T = 10$, $\gamma = 2$ $\mu_d = \mu_n = 0.15$, $r = 0.10$, $\sigma = 0.20$, $\Delta_d = 6.5$ hours, $\Delta_n = 17.5$ hours, $\alpha = \theta = 0.01$.

### 4.4.3 Intraday trading volume

It is well known that the daily trading volume is U-shaped, i.e, the trading volumes at market open and market close are much higher than the rest of a day. Our model predicts such a trading pattern.

Figure 4.4 displays the fraction of total buying and selling volume that occurred within a given time interval against time. It shows that an investor trades much more at the open and at the close than during other time of a trading period. This is because investors cannot trade overnight and thus it is optimal to adjust his portfolio before market closes. Since there is no trading overnight, the position may be out of the no-transaction region by the next market open, therefore they may also trade more to main the position within no-transaction region at market open. So this rationale leads to the U-shaped trading volume. Moreover, when the overnight volatility is small, stock would be more attractive to investors in nighttime compared to during daytime. So it is optimal to hold more stock overnight and reduce the stock position during daytime. In this way, investors typically buy more at market close and sell more at market open, as what we see in Figure 4.4.

Chapter 5

# Conclusion

This thesis contains a two-fold study on the portfolio selection problem for a CRRA investor who faces proportional transaction costs and finite investment horizon. Two factors are considered separately: consumption, and market closure with time-varying stock return dynamics.

## 5.1 Optimal investment with consumption

For the optimal investment problem with consumption, mathematically speaking, it is formulated as a singular stochastic control problem, with the trading policy and consumption strategy as controls. The optimization objective is to maximize CRRA utility from both terminal wealth and cumulative consumption. Then in terms of the Hamilton-Jocabi-Bellman equation, we obtain a degenerate parabolic variational inequality with gradient constraints on the value function (denoted as $w(x, \tau)$), which gives rise to two free boundaries.

Since it is not straightforward to solve the variational inequality with gradient constraints directly, we formally take partial derivative in the original variational inequality and then arrive at a standard variational inequality (i.e. an obstacle problem) that some partial derivative of the value function (denoted as $v(x, \tau)$) satisfies. This approach is the same as in Dai and Yi (2009), or as in Dai, Jiang and Yi (2007). Once regularity of function $v(x,\tau)$ is obtained, regularity of the original value function $w(x,\tau)$ can be established by showing equivalence between the original variational inequality with gradient constraints and the double obstacle problem. And, this equivalence can be proved given that the free boundary comes forth from the double obstacle problem is smooth. So finally, the problem reduces to show the smoothness of free boundary. Dai and Yi (2009) followed the arguments in Friedman (1975) to prove this smoothness and their argument relied on the monotonicity in time of function $v(x,\tau)$ which asserted $v_\tau(x,\tau) \leq 0$.

For our problem in this thesis, due to the presence of consumption, the double obstacle problem obtained itself is not a self-contained system. In the differential operator, one item which contains function $w(x,\tau)$ from the original variational inequality with gradient constraints is involved. It is this extra term resulted from consumption that leads to the most pivotal difficulty of this topic.

Dai, Jiang and Yi (2007) attempted to use the same approach as in Dai and Yi (2009) to attack this problem. The argument about equivalence was much more complicated than that in Dai and Yi (2009) due to the non-self-contained property of the double obstacle problem. Fortunately, Schauder's fixed point theorem could be employed to conquer this difficulty, again, provided that the free boundary is smooth. In Dai, Jiang and Yi (2007), they followed Friedman (1975) to prove the smoothness. Their arguments was based on the monotonicity in time of function $v(x,\tau)$, i. e. $v(x,\tau) \leq 0$; while this monotonicity in time was assured only when $\gamma < 1$ and $\beta < (1-\gamma)r$.[1] To avoid imposing such technical conditions, we aim at proving the smoothness of free boundary bypassing monotonicity in time of $v(x,\tau)$. Dai, Xu and Zhou (2008) set up a template for us. Making use of the bootstrap technique, we obtain the smoothness of free boundary by showing the cone property in the problem.

Now let us go over the logic path of the arguments more specifically.

To obtain regularity of solution to the double obstacle problem, we first study the

[1] In Dai and Yi (2009) and Dai, Jiang and Yi (2007), their "$\gamma$" was equivalent to the "$(1-\gamma)$" in this thesis, as mentioned before.

double obstacle problem given a known function $w_{\text{known}}(x, \tau)$ (independent from the original variational with gradient constraints.) with certain prescribed properties. In this way, we obtain the existence and smoothness of function $v(x, \tau)$, which is dependent on $w_{\text{known}}(x, \tau)$. Moreover, we proved the existence of two free boundaries ($x_{s,w_{\text{known}}}(\tau)$ and $x_{b,w_{\text{known}}}(\tau)$) representing the optimal selling and buying strategies. Most importantly, we obtain the infinite smoothness of the trading boundary ($x_{s,w_{\text{known}}}(\tau) \in C^{\infty}$) by means of bootstrap technique. The smoothness of the free boundaries prepares a sound foundation for later argument to retrieve regularity of the original variational inequality with gradient constraints.

Next, in terms of Schauder's fixed point theorem, we manage to show that the original variational inequality with gradient constraints on value function $w(x, \tau)$ combined with the double obstacle problem on the partial derivative function $v(x, \tau)$ uniquely share a solution triple $(w(x, \tau),\ v(x, \tau)\ x_s(\tau))$. In this solution triple, $w(x, \tau)$ is the value function in the original variational inequality, $v(x, \tau)$ is the solution to the double obstacle problem with $w_{\text{known}}(x, \tau) = w(x, \tau)$, and $x_s(\tau)$ is the corresponding free boundary.

In Dai and Yi (2009), the properties of free boundaries from optimal investment problem without consumption has been fully characterized. Based on their results and a comparative proposition, we are finally able to analyze the behaviors of the free boundaries (optimal trading strategies) in our model with consumption. Compared with the no-consumption case, the free boundaries are no longer monotone, while most other properties remain valid. For instance, there is a critical time after which it is never optimal to purchase stocks. The no-trading region is always in the first quadrant if and only if $\mu - r - \gamma\sigma^2 \leq 0$, which means that leverage is always suboptimal if risk premium is non-positive.

For the portfolio selection problem with consumption, finally, we would like to mention that our approach relies on the connection between singular control and optimal stopping, which is well known in the field of singular stochastic control, but has never been revealed for the present problem. This approach can also be utilized to handle the infinite horizon problems.

## 5.2 Optimal investment with market closure

Then we consider the optimal investment problem with market closure and time-varying return dynamics, where consumption is absent. In this case, we show that incorporating the well-established return dynamics across trading and nontrading periods *alone* can generate more than a first order effect of transaction costs on asset pricing. In addition, we find that adopting strategies prescribed by standard portfolio selection models that assume a continuously open market (e.g., Merton (1987)) can result in significant utility loss. Furthermore, consistent with empirical evidence, our model predicts that trading volumes at market close and market open are much larger than the rest of trading times.

Specifically, we consider a continuous-time optimal portfolio selection problem of an investor with a finite horizon who can trade a risk-free asset and a risky asset. He faces proportional transaction costs in trading the stock. Different from the standard literature and consistent with empirical evidence, we assume market closes periodically and stock return volatilities differ across trading and nontrading periods. We show the existence, uniqueness, and smoothness of the optimal trading strategy. We also explicitly characterize the solution to the investor's problem and derive certain helpful comparative statics on the optimal trading strategies. Our extensive numerical analysis, using parameter estimates used by Constantinides (1986), demonstrates that in contrast to the standard conclusion that transaction costs only have a *second-order* effect, transaction costs can have a more than *first-order* effect if one takes into account the time varying volatilities across trading and nontrading periods. In particular, the liquidity premium to transaction cost (LPTC) ratio could be well above one. Indeed, the LPTC ratio can be more than 20 times higher than what Constantinides finds for reasonable parameter values.

An intuitive explanation for higher liquidity premium in the presence of time-varying return dynamics is that when return dynamics varies across time, investors tend to trade more often in this certain circumstance to adjust their positions, and thus incur more transaction cost payments. Surprisingly, we show that the real reason contradicts our intuition. Between the two sources of liquidity premium from transaction costs, direct transaction cost payment and the relative suboptimal trading strategy in the absence of transaction costs, the suboptimality of strategies dominates in our model. As a consequence, investors in our model trade much *less* frequently but with larger average trading size than those in Constantinides' model. This is because with the large discrepancy between volatilities across trading and non-trading periods, investors are "forced" to widen the no-transaction region significantly to avoid paying too much transaction costs from trading frequently and consequently their stock position is much further from the allocation that is optimal in the absence of transaction costs. Although investors in our model still pay more than double the transaction costs than those in Constantinides' model, it is essentially this substantial suboptimality of the trading strategy that produces the high liquidity premium in our model.

We also show that the "optimal" trading strategy prescribed by the standard portfolio selection literature can result in large utility loss. For example, given constant relative risk aversion (CRRA) preferences and constant investment opportunity set, the optimal trading strategy is to keep a constant fraction of wealth in the stock in the absence of transaction costs. We show that implementing this strategy in a market with market closure and time-varying volatilities can cost as much as 12.29% of initial wealth for an investor with risk aversion coefficient of 2 and investment horizon of 10 years. Intuitively, assuming a constant volatility results in overinvestment or underinvestment almost all the time, thus causes substantial utility loss.

Finally, periodic market closure and volatility difference across trading and non-trading periods would imply a U-shaped trading volume pattern, which means trading volume at market open and close can be much higher than other trading times due to discrete position adjustments. This trading volume patter is strongly supported by empirical evidence.

To conclude, this thesis has investigated finite horizon portfolio selection problem with consumption *or* with market closure accompanied by time-varying stock return dynamics. The portfolio selection problem with *both* consumption and market closure remains unclear yet. The difficulty lies in how to understand consumption and prescribe appropriate boundary conditions on the line of $x = 0$ during market closure. We leave it as a future research topic.

# Appendix A

**Lemma A.0.1.** *(Dai and Yang (2009)'s work) Let $v_\delta(x,\tau)$ be the solution to problem (3.19) in which $w(x,\tau)$ satisfies (3.12)-(3.15). Then there is a positive constant $K$ independent of $\delta$ and $R$, such that*

$$-\frac{K}{(x+1-\alpha)^2} \leq (v_\delta)_x\,. \tag{A.1}$$

*Proof.* Since it has been proved when $\gamma < 1$, we only need to provide a proof in the case of $\gamma > 1$. In this case, $(1-\gamma) < 0$. Instead of considering $(v_\delta)_x$, we consider the following quantity $(v_\delta)_x + (1-\gamma)v_\delta^2$,which is inspired by the change of optimal consumption w.r.t. dollar value in bank account.

We aim at proving that there exists a $K$, such that

$$(v_\delta)_x + (1-\gamma)v_\delta^2 \geq -\frac{K}{(x+1-\alpha)^2}.$$

and note that $(1-\gamma) < 0$, thus (A.1) will follow naturally.

Again without loss of generality, we can confine ourselves to the region

$$\mathcal{M} \equiv \left\{(x,t) \in \Omega_T^R : \frac{1}{x+1+\theta} < v_\delta < \frac{1}{x+1-\alpha}\right\}.$$

Following the notations in proving Proposition (1), we denote $p = \partial_x v_\delta$ and $q = v_\delta^2(x,t)$. we already have

$$p_\tau - \mathcal{L}^* p + (e^{\gamma w} v_\delta)^{-\frac{1}{\gamma}}(p_x + q_x) - \frac{1}{\gamma}\left(e^{(1-\gamma)w} v_\delta\right)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w}((1-\gamma)\,v_\delta w_x + p)(p+q)$$
$$-4\,(1-\gamma)\,\sigma^2 x v_\delta p - (1-\gamma)\,\sigma^2 x^2 v_\delta p_x = (1-\gamma)\,\sigma^2 x^2 p^2 + (1-\gamma)\,\sigma^2 q, \quad \text{in } \mathcal{M},$$

and

$$\begin{aligned} & q_\tau - \mathcal{L}^* q + 2v_\delta (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} (p+q) - 2(1-\gamma)\sigma^2 x^2 pq \\ = \;& -\sigma^2 (x q_x + q) - \sigma^2 x^2 p^2 + 2(1-\gamma)\sigma^2 x v_\delta q - \delta\sigma^2 p^2, \quad \text{in } \mathcal{M}, \end{aligned}$$

where

$$\mathcal{L}^* p = \left(\frac{1}{2}\sigma^2 x^2 + \delta\right) p_{xx} - (\mu - r - (2+\gamma)\sigma^2) x p_x - (2\mu - 2r - (1+2\gamma)\sigma^2) p.$$

Let $H = p + (1-\gamma) q$, then $H$ satisfies

$$\begin{aligned} & H_\tau - \mathcal{L}^* H + (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} H_x \\ & -\frac{1}{\gamma} (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w} [\, H^2 + ((1-\gamma) v_\delta w_x + q - 2(1-\gamma) q) H \,] \qquad \text{(A.2)} \\ & -4(1-\gamma)\sigma^2 x v_\delta H - (1-\gamma)\sigma^2 x^2 v_\delta H_x + 2(1-\gamma) v_\delta (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} H - 2(1-\gamma)^2 \sigma^2 x^2 q H \\ = \;& -\gamma (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} q_x + (1-\gamma)(e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w} q (v_\delta w_x - q) \\ & -4(1-\gamma)^2 \sigma^2 x v_\delta q - (1-\gamma)^2 \sigma^2 x^2 v_\delta q_x - 2\gamma(\gamma-1) v_\delta (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} q - 2(1-\gamma)^3 \sigma^2 x^2 q^2 \\ & +(1-\gamma)\sigma^2 q - (1-\gamma)\sigma^2 (x q_x + q) + 2(1-\gamma)^2 \sigma^2 x v_\delta q - (1-\gamma)\delta\sigma^2 p^2 \\ \geq \;& (1-\gamma)(e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w} q v_\delta w_x - 2(1-\gamma)^2 \sigma^2 x v_\delta q - (1-\gamma)\sigma^2 x q_x \\ = \;& (1-\gamma)(e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} v_\delta^2 w_x - 2(1-\gamma)\sigma^2 x v_\delta H, \qquad \text{(A.3)} \end{aligned}$$

where we have used $q_x = 2v_\delta (v_\delta)_x < 0$. Define a new differential operator

$$\begin{aligned} \mathcal{T} H \;\equiv\; & \mathcal{L}^* H - \left[ (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} - (1-\gamma)\sigma^2 x^2 v_\delta \right] H_x \\ & - \left[ \frac{1}{\gamma} (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} ((4(1-\gamma) - 1 - 2(1-\gamma)^2) v_\delta - (1-\gamma) w_x) \right. \\ & \qquad \left. -2(1-\gamma)\sigma^2 x v_\delta - 2(1-\gamma)^2 \sigma^2 x^2 v_\delta^2 \right] H \\ & + (1-\gamma)(e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}} v_\delta^2 w_x. \end{aligned}$$

It follows from (A.3) that

$$H_\tau - \mathcal{T} H \geq \frac{1}{\gamma} (e^{(1-\gamma)w} v_\delta)^{-\frac{1}{\gamma}-1} e^{(1-\gamma)w} H^2 \geq 0.$$

Let $W = -e^{K_1\tau}(x+1-\alpha)^{-2}$, where $K_1$ is sufficiently large and independent of $\delta$ and $R$. Noticing $w$ satisfies (3.12)-(3.15), it is not hard to get

$$W_\tau - \mathcal{T}W \leq \frac{-K_1 e^{K_1\tau}}{(x+1-\alpha)^2} + \frac{\hat{K}_1 e^{K_1\tau}(x^2+1)}{(x+1-\alpha)^4} \leq 0,$$

where $\hat{K}_1$ is also a constant. Clearly $H \geq W$ on the boundary of $\mathcal{M}$. By comparison principle, we then obtain $H \geq W$ in $\mathcal{M}$, which yields the desired result by taking $K = e^{K_1 T}$. The proof is complete. $\square$

# Bibliography


[1] Akian, M., J. L. Menaldi, and A. Sulem (1996): On an investment-consumption model with transaction costs, *SIAM Journal of Control and Optimization*, 34, 329-364.

[2] Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics*, 17, 223C249.

[3] Bertsimas, D., and Lo A. W. (1998): Optimal control of execution costs, *Journal of Financial Markets*, 1(1998), 1-50.

[4] Brennan, Michael, T. Chordia, and A. Subrahmanyam (1998): Alternative factor specifications, security characteristics, and the cross section of expected returns, *Journal of Financial Economics* 49, 345–374.

[5] Cocco, J. F. (2005): Portfolio choice in the presence of housing, *Review of Financial Studies*, 18(2), 535-567.

[6] Constantinides, G. M. (1986): Capital market equilibrium with transaction costs, *The Journal of Political Economy* 94, 842–862.

[7] Cvitanic, J., and I. Karatzas (1996): Hedging and portfolio optimization under transaction costs: a martingale approach, *Mathematical Finance*, 6(2), 133-165.

[8] Dai, M., L. S. Jiang and F. H. Yi (2007): Analysis of a degenerate parabolic variational inequality with gradient constraints from optimal investment and consumption with transaction costs, available at *SSRN: http://ssrn.com/abstract=955465*

[9] Dai, M., L. S. Jiang, P. F. Li and F. H. Yi (2009): Finite horizon optimal investment and consumption with transaction costs, *SIAM Journal on Control and Optimization*, (2009), 48(2), 1134-1154.

[10] Dai, M., P. F. Li and H. Liu (2009): Market closure, portfolio selection and Liquidity Premia, *working paper.*

[11] Dai, M., Z. Q. Xu and X. Y. Zhou (2008): Continuous-time mean-variance portfolio selection with proportional transaction costs, *working paper.*

[12] Dai, M., and Z. Yang (2009): A note on finite horizon optimal investment and consumption with transaction costs, *working paper.*

[13] Dai, M., and F. H. Yi (2009): Finite horizon optimal investment with transaction costs: a paraoblic double obstacle problem, *Journal of Differential Equation* , 246, 1445-1469.

[14] Dai, M., and Y. F. Zhong (2008): Penalty method for continuous-time portfolio selection with proportional transaction costs, accepted for publication in *Journal of Computational Finance.*

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

[16] Davis, M. H. A., V. G. Panas, and T. Zariphopoulou (1993): European option pricing with transaction costs, *SIAM Journal of Control and Optimization* , 31(2), 470-493.

[17] Eleswarapu, V. R. (1997): Cost of transacting and expected returns in the Nasdaq market, *Journal of Finance*, 52, 2113-2127.

[18] Evans, L. C. (1979): A second-order elliptic equation with gradient constraint, *Communication in Partial Differential Equations* , 4(5), 555-572.

[19] Fleming, W. H., and H. M. Soner (2006): *Controlled Markov Processes and Viscosity Solutions*, 2nd edition, Springer-Verlag, New York.

[20] French, K. R., R. Roll (1986): Stock return variances: the arrival of information and the reaction of traders. *Journal of Financial Economics* 17, 5-26.

[21] Friedman, A. (1975): Parabolic variational inequality in one space dimension and smoothness of the free boundary, *Journal of Functional Analysis*, 18(3), 151-176.

[22] Friedman, A. (1982): *Variational principles and free-boundary problems*, Wiley, New York.

[23] Gomes, F., and A. Michaelides (2005): Optimal life-cycle asset allocation: understanding the empirical evidence, *Journal of Finance*, 60(2), 869 - 904.

[24] Hu, B. (1986): Fully nonlinear elliptic equations with gradient constraint, *Beijing Daxue Xuebao*, 5, 78-91.

[25] Ishii, H. and S. Koike (1983): Boundary regularity and uniqueness for an elliptic equation with gradient constraint, *Communication in Partial Differential Equations*, 4, 317-346.

[26] Janecek, K., and S. E. Shreve (2004): Asymptotic analysis for optimal investment and consumption with transaction costs, *Finance & Stochastics*, 8, 181-206.

[27] Jang, B. G., H. K. Koo, H. Liu, and M. Loewenstein (2007): Liquidity premia and transaction costs, *Journal of Finance*, 62, 2329C2366.

[28] Karatzas, I., and S. E. Shreve (1984): Connections between optimal stopping and singular stochastic control i: monotone follower problems, *SIAM J. Control and Optimization*, 22(6), 856-877.

[29] Karatzas, I., and S. E. Shreve (1998): *Methods of Mathematical Finance.* Springer-Verlag, New York.

[30] Lai, T. L., and T.W. Lim (2005): Singular stochastic control in optimal investment and hedging in the presence of transaction costs, *Probability, Statistics, and Their Applications, IMS Lecture Notes Monogr. Ser. 41, Inst. Math. Statist., Beachwood, OH*, pp. 209-227.

[31] Liu, H. (2004), Optimal consumption and investment with transaction costs and multiple risky assets, *Journal of Finance,* 59, 289–338.

[32] Liu, H., and M. Loewenstein (2002): Optimal portfolio selection with transaction costs and finite horizons, *Review of Financial Studies*, 15(3), 805-835.

[33] Lockwood, Larry J. and Scott C. Linn (1990): An examination of stock market return volatility during overnight and intraday periods, 1964–1989, *Journal of Finance,* 45, 591–601.

[34] Magill, M. J. P., and G. M. Constantinides (1976): Portfolio selection with transaction costs, *Journal of Economic Theory*, 13, 264-271.

[35] Markowitz, H. M. (1952): Portfolio selection, *Journal of Finance*, 7(1), 77-91.

[36] Merton, R. C. (1969): Lifetime portfolio selection under uncertainty: the continuous-time case, *Review of Economics and Statistics*, 51(3), 247-57.

[37] Merton, R. C. (1971): Optimal consumption and portfolio rules in a continuous time model, *Journal of Economic Theory*, 3, 373-413.

[38] Merton, R. C. (1987): A simple model of capital market equilibrium with incomplete information, *Journal of Finance*, 42, 483-510.

[39] Shreve, S. E. and H. M. Soner (1994): Optimal investment and consumption with transaction costs, *Annals of Applied Probability*, 4, 609-692.

[40] Shreve, S. E., H. M. Soner, and G.-L. Xu (1991): Optimal investment and consumption with two bonds and transaction costs, *Mathematical Finance*, 1(3), 53-84.

[41] Soner, H. M. and S.E. Shreve (1991): A free boundary problem related to singular stochastic control: the parabolic case, *Communication in Partial Differential Equations*, 16(2&3), 373-424.

[42] Stoll, H. R., and R. E. Whaley (1990), Stock market structure and volatility, *Review of Financial Studies* 3, 37–41.

[43] Tsiakas, I. (2008), Overnight information and stochastic volatility: A study of European and US stock exchanges, *Journal of Banking and Finance,* 32, 251-268.

[44] Vayanos, D. (1998): Transaction costs and asset prices: A dynamic equilibrium model, *Review of Financial Studies,* 11, 1–58.

[45] Wiegner, M. (1981): The $C^{1,1}$-character of solutions of second order elliptic equations with gradient constraint, *Communication in Partial Differential Equations*, 6(3), 361-371.

[46] Yong, J. M., and X. Y. Zhou (1999): *Stochastic Controls: Hamiltonian Systems and HJB Equations,* Springer-Verlag, New York.

[47] Zhu, H. (1992): Generalized solution in singular stochastic control: the nondegenerate problem, *Applied Mathematics and Optimization*, 25, 225-245.

**FINITE HORIZON PORTFOLIO SELECTION WITH TRANSACTION COSTS**

**LI PEIFAN**

**2009**