

Problems in Mathematical Finance Related to Transaction Costs and Model Uncertainty

by

Yuchong Zhang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied and Interdisciplinary Mathematics)
in the University of Michigan
2015

Doctoral Committee:

Professor Erhan Bayraktar, Co-Chair
Professor Uday Rajan, Co-Chair
Professor Joseph G. Conlon
Assistant Professor Sergey Nadtochiy
Professor Virginia R. Young

© Yuchong Zhang 2015
All Rights Reserved

To my mom and dad

ACKNOWLEDGEMENTS

I can never overstate my gratitude to my advisor Professor Erhan Bayraktar, who introduced me to the field of mathematical finance, and has trained me in every aspect of being an applied mathematician, from doing research to giving talks, writing papers and referee reports, and interacting with the wider academic community. Just about three years ago, I was still a confused student who would cry in his office for not knowing how to do research. Without his “tough” requirement and at the same time, enormous guidance and encouragement, this thesis would not have been possible. I am also grateful for his financial support in the past two winter semesters. In addition, I want to thank him for organizing the Financial/Actuarial Mathematics Seminar from which I greatly benefited.

I would like to thank Professors Uday Rajan, Sergey Nadtochiy, Joseph Conlon and Virginia Young for serving on my committee and reading my thesis. In particular, I deeply appreciate the interest Professor Virginia Young has shown in my work, and her valuable feedbacks on the lifetime ruin problem.

Many other professors have helped in my Ph.D. study, thus indirectly contributed to the writing of this thesis. I want to thank Professor Mattias Jonsson for being supportive and for twice having me as a teaching assistant. I owe many thanks to Professor Peter Miller who provided a lot of assistance in my first year here at Michigan. I am also greatly indebted to my undergraduate teacher Professor Zhouping Xin whose encouragement helped me survive some of the hard times in

graduate school.

Summer financial support from the Department of Mathematics and Rackham Graduate School is acknowledged. I am also grateful for all the travel grants I received from Rackham Graduate School, the Society of Industrial and Applied Mathematics, and the organizers of the 7th European Summer School in Financial Mathematics.

I would like to express my appreciation to my friend and collaborator Zhou Zhou who contributed to the discussion in Section 2.2 of this thesis. His diligence and productiveness is a constant source of motivation for me. My appreciation also goes to Dr. Xiang Yu who discussed mathematics with me regularly in my third year; to Xueying Hu and Yu-Jui Huang, for all the advices they gave; and to my fellow math students Xiaolei Zhao, Tengren Zhang, Purvi Gupta, Zhibek Kadyrsizova, and many other friends, for their company in this otherwise lonely journey. Special thanks to my tennis-mates in Ann Arbor Chinese Tennis Club for making my life colorful and allowing me to rejuvenate myself on court every week.

Finally, I want to thank my father, Zhengjun Zhang, and my mother, Yu Hu, for their unconditional love and support.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	vii
ABSTRACT	viii
 CHAPTER	
I. Introduction	1
 II. Fundamental Theorem of Asset Pricing Under Transaction Costs and Model Uncertainty	
2.1 Introduction	5
2.1.1 The uncertainty set \mathcal{P}	9
2.2 When bid-ask spread is on hedging options only	10
2.2.1 Fundamental theorem with robust no-arbitrage	10
2.2.2 A sharper fundamental theorem with a non-redundancy assumption	17
2.3 When bid-ask spread is on one dynamically traded asset	21
2.3.1 The market model and main results	21
2.3.2 The building block: the one-period case	25
2.3.3 The multi-period case	29
2.3.4 Existence of an optimal superhedging strategy	42
2.4 Further discussion	46
2.4.1 Challenges with multi-asset extension	46
2.4.2 Challenges with superhedging duality	47
 III. Minimizing the Probability of Lifetime Ruin Under Ambi- guity Aversion	
3.1 Introduction	49
3.2 Problem formulation	54

3.3	Explicit solution in the zero-hazard rate case	62
3.4	Existence and uniqueness of viscosity solution	65
3.4.1	Comparison principle	65
3.4.2	Perron’s method	68
3.5	Regularity	71
3.5.1	An equivalent convex problem	71
3.5.2	From convexity to C^∞	72
3.6	Verification	78
3.6.1	Regularity of π^*	78
3.6.2	Verification theorem	81
3.6.3	Proof of Theorem 3.2.4.	87
3.7	Properties of the value function and the optimal investment policy	88
3.7.1	When does the value function lose convexity?	88
3.7.2	Reaching the safe level or not?	91
3.8	Asymptotic expansion for small ε	92
3.9	Numerical examples	95
3.9.1	Should the individual care about robustness?	99

**IV. Stochastic Perron’s Method for the Lifetime Ruin Problem
Under Transaction Costs 101**

4.1	Introduction	101
4.2	Problem formulation	104
4.2.1	Upper and lower bounds on the value function	106
4.2.2	Random initial condition and admissible controls	108
4.2.3	Main results	111
4.3	Stochastic supersolution	111
4.4	Stochastic subsolution	122
4.5	Comparison principle	131
4.5.1	Proof of Theorem 4.2.5.	136

APPENDIX 138

BIBLIOGRAPHY 143

LIST OF FIGURES

Figure

3.1	Robust ruin probabilities and distorted Sharpe ratios when $\lambda = 0$. . .	64
3.2	Robust ruin probabilities.	96
3.3	Cole-Hopf transform of three non-convex curves in Figure 3.2. . . .	96
3.4	Optimal investments.	97
3.5	Distorted Sharpe ratios.	97
3.6	Optimal investments with different ruin levels.	98

ABSTRACT

This thesis is devoted to the study of three problems in mathematical finance which involve either transaction costs or model uncertainty or both.

In Chapter II, we investigate the Fundamental Theorem of Asset Pricing under both transaction costs and model uncertainty, where model uncertainty is described by a family of probability measures, possibly non-dominated. We first show that the FTAP and super-hedging results of [25] can be extended to the case in which only the options available for static hedging (hedging options) are quoted with bid-ask spreads. In this set-up, we need to work with the notion of robust no-arbitrage which turns out to be equivalent to no-arbitrage under the additional assumption that hedging options with non-zero spread are non-redundant. A key result is the closedness of the set of attainable claims, which requires a new proof in our setting. Next, we look at the more difficulty case when the market consists of a money market and a dynamically traded stock with bid-ask spread. Under a continuity assumption, we prove using a backward-forward scheme that the absence of arbitrage in a quasi-sure sense is equivalent to the existence of a suitable family of consistent price systems.

In Chapter III, we study the problem where an individual targets at a given consumption rate, and seeks to minimize the probability of lifetime ruin when she does not have perfect confidence in the drift of the risky asset. Using stochastic control, we characterize the value function as the unique classical solution of an associated

Hamilton-Jacobi-Bellman (HJB) equation, obtain feedback forms for the optimal investment and drift distortion, and discuss their dependence on various model parameters. In analyzing the HJB equation, we establish the existence and uniqueness of viscosity solution using Perron's method, and then upgrade regularity by working with an equivalent convex problem obtained via the Cole-Hopf transformation. We show the original value function may lose convexity and the Isaacs condition may fail. Numerical examples are also included to illustrate our results.

In Chapter IV, we adapt stochastic Perron's method to the lifetime ruin problem under proportional transaction costs which can be formulated as a singular stochastic control problem. Without relying on DPP, we characterize the value function as the unique viscosity solution of an associated HJB variational inequality. We also provide a complete proof of the comparison principle which is the main assumption of stochastic Perron's method.

CHAPTER I

Introduction

This thesis is the study of three different problems in mathematical finance which share the common ingredients of transaction costs or model uncertainty. The first problem, which we examine in Chapter II, is concerned with the Fundamental Theorem of Asset Pricing under both proportional transaction costs and model uncertainty. The second and the third problems, discussed in Chapters III and IV, respectively, fall into the field of robust or singular stochastic control. The importance of transaction costs (or market friction in general) and model uncertainty not only lies in that they are more realistic modelings of the financial market, but also in that they require sophisticated mathematical treatments different from their frictionless or fixed-model counterparts in each of the problems we study.

The Fundamental Theorem of Asset Pricing (FTAP), as suggested by its name, is one of the most important theorems in mathematical finance and has been established in many different settings: discrete and continuous, with and without transaction cost. It relates no-arbitrage concepts to the existence of certain fair pricing mechanisms, which provides the rationale for why in duality theory, it is often reasonable to assume that the dual domain is non-empty. In Chapter II, we study the FTAP for a discrete-time, finite horizon market where trading is subject to propor-

tional transaction costs and the price dynamics is modeled by a family of probability measures, usually non-dominated. The theories with only transaction costs or only model uncertainty are established in [56], [39], [71] and [25], [1], respectively. We attempt to fill in the gap when both market friction and model uncertainty are present, and when we are in the quasi-sure framework. Generalizing to the non-dominated case means the usual separation argument does not work and we have to resort to a local analysis followed by a pasting argument as in [25]. Adding transaction costs on top of model uncertainty brings in the additional difficulty that the absence of arbitrage in a multi-period market is not equivalent to the absence of arbitrage in all single-period markets. As a consequence, the pasting of single-period dual elements can be problematic. We design a backward-forward scheme to address this issue. In particular, the backward part corresponds to replacing the original market by a new market which embeds future price information into the current prices. This chapter is based on [17] and [20]. Part of the work has been presented in the Financial/Actuarial Mathematics Seminar at the University of Michigan (September 4, 2013), and the Labex Louis Bachelier - SIAM-SMAI Conference on Financial Mathematics at University Paris Diderot (June 20, 2014).

The problem of how an individual, usually a retiree, with a target consumption rate, should invest in a risky financial market to minimize the probability of lifetime ruin was first studied in [82] and later in a series of extensions [14], [15], [16], [7], all in a fixed-model setting. In reality, there may be some good estimates of the price volatility, but drift estimation is almost impossible; it would require centuries of data to obtain a reliable estimate. A natural approach is to extract from the available data a reference model, and penalize other models based on their deviation from the reference model. How hard to penalize depends on how averse the agent is to model

uncertainty, also called ambiguity or Knightian uncertainty. This leads to a robust lifetime ruin problem which we solve in Chapter III using the theory of stochastic control and viscosity solutions. In analyzing the associated Hamilton-Jacobi-Bellman (HJB) equation, we do not rely on the dynamic programming principle (DPP) which is very complicated when the optimization problem resembles a stochastic differential game. Instead, we use Perron's method to obtain the existence and uniqueness of a viscosity solution, and then upgrade regularity by working with an equivalent convex problem obtained through the Cole-Hopf transformation. Chapter III is based on [18]. Part of this work has been presented in the Financial/Actuarial Mathematics Seminar at the University of Michigan (February 19, 2014), the 2014 SIAM Conference on Financial Mathematics and Engineering (November 15, 2014), the Statistics and Actuarial Science Department Seminar at the University of Waterloo (January 26, 2015) and the ORFE Colloquium at Princeton University (February 12, 2015).

Stochastic Perron's method is introduced in [9], [10] and [11] as a way to obtain a PDE characterization of the value function of a stochastic control problem, without relying on the DPP. It is a direct verification approach in that it first constructs a solution to the HJB equation, and then verifies such a solution is the value function. But unlike the classical verification, it does not require regularity; uniqueness acts as a substitute for verification. The method has been applied to linear problems, Dynkin games, HJB equations for regular control problems, (regular) exit time problems and zero-sum differential games. In Chapters IV, we adapt the method to another type of problems: singular control problems. In particular, we focus on the specific problem of minimizing the probability of lifetime ruin when buying and selling stocks incurs proportional transaction costs, and demonstrate how to take care of jumps and gradient constraints in the adaption of stochastic Perron's method. Chap-

ter IV is based on [19]. Part of the work has been presented in the 7th European Summer School in Financial Mathematics at the University of Oxford (September 4, 2014), the Financial/Actuarial Mathematics Seminar at the University of Michigan (November 5, 2014), and the IMS PDE Seminar at the Chinese University of Hong Kong (December 18, 2014).

CHAPTER II

Fundamental Theorem of Asset Pricing Under Transaction Costs and Model Uncertainty

2.1 Introduction

In this chapter, we investigate the FTAP and in some cases, the super-hedging theorem, for a discrete time, finite horizon financial market under both proportional transaction costs and model uncertainty. When the market is frictionless and modeled by a single probability measure, the classical result by Dalang-Morton-Willinger [30] asserts there is no-arbitrage if and only if there exists a martingale measure. With proportional transaction cost, martingale measure is replaced by the concept of consistent price system (CPS) or strictly consistent price system (SCPS). Equivalence between no-arbitrage and existence of a CPS is established by Kabanov and Stricker [56] for finite probability space Ω , and by Grigoriev [39] when the dimension is two. Such equivalence in general does not hold in higher dimensions and when Ω is infinite (see Section 3 of [71] and page 128-129 of [55] for counter examples). For such an equivalence one needs the notion of robust no-arbitrage introduced by Schachermayer [71], where he showed it is equivalent to the existence of an SCPS. Alternatively, robust no-arbitrage can be replaced by strict no-arbitrage plus efficient friction (see Kabanov et al. [54]). There exist a few different proofs of the FTAP un-

der transaction costs. Besides the proof in [71] and [54] which rely on the closedness of the set of hedgeable claims and a separation argument, there is a utility-based proof by Smaga [75] and proofs based on random sets by Rokhlin [66].

In recent years, model uncertainty has gained a lot of interest since it corresponds to a more realistic modeling of the financial market. By model uncertainty, we mean a convex family \mathcal{P} of probability measures, usually non-dominated. Each member of \mathcal{P} represents a possible model for the asset price behavior. One should think of \mathcal{P} as being obtained from calibration to the market data. We have a collection of measures rather than a single one because we do not have point estimates but confidence intervals. The non-dominated case is generally much harder because the classical separation argument used to construct the dual element does not work. Without transaction costs, the recent work by Bouchard and Nutz [25] used a local analysis to establish the equivalence between the absence of arbitrage in a quasi-sure sense and the existence of an “equivalent” family of martingale measures. Acciaio et al. [1] obtained a different version of the FTAP by working with a different no-arbitrage condition which excludes model-independent arbitrage, i.e. arbitrage in a sure sense only. The two different frameworks are often referred to as the “quasi-sure” framework and the “model-free” framework, respectively. The former correspond to \mathcal{P} being an arbitrary convex collection of probability measures, and the later corresponds to \mathcal{P} being the collection of all probability measures.

We attempt to generalize some results on FTAP with proportional transaction costs from the case when there is a dominating measure to the case when a dominating measure may be absent, under the quasi-sure framework. On a related note, Dolinsky and Soner [35] proved a frictional super-hedging theorem (by first discretizing the state space and then taking a limit) and stated the FTAP as a corollary in

the model-free framework. We follow a different methodology, and are able to work with a more general structure on the proportional transaction costs instead of taking it a constant, which is useful as this proportion in real markets is likely to change with changing market conditions over time.

We begin in Section 2.2 with a simpler setting where the financial market consists of a zero-interest money market, some dynamically traded stocks and some statically traded options. We assume that stocks are liquid and trading in them does not incur transaction costs, but that the options are less liquid and their prices are quoted with a bid-ask spread. Our first goal is to obtain a criteria for deciding whether the collection of models represented by \mathcal{P} is viable or not. Given that \mathcal{P} is viable we would like to obtain the range of prices for other options written on the stocks. The dual elements in these result are martingale measures that price the hedging options correctly (i.e., consistent with the quoted prices). As in classical transaction costs literature, we need to replace the no-arbitrage condition by the stronger *robust no-arbitrage condition* which we later prove to be equivalent to no-arbitrage under an additional assumption that hedging options with non-zero spread are *non-redundant*, i.e. not replicable by other hedging options. As a result, we obtain two versions of the FTAP and super-hedging theorem, one with robust no-arbitrage (Section 2.2.1), the other with the non-redundancy assumption (Section 2.2.2).

In Section 2.3, we consider the more difficult problem when dynamic trading incurs transaction costs, but restrict ourselves to the case when there is a single stock and no hedging options. In the absence of a dominating measure, the main idea, as initiated in [25], is to proceed in a local fashion: first obtain dual elements for each single-period and then do pasting using a suitable measurable selection theorem. However, the multi-period case turns out to be quite different when transaction

costs are added. A distinct feature for frictionless markets is that the absence of arbitrage for the multi-period market is equivalent to the absence of arbitrage in all single-period markets. So it is enough to look at each single period separately and paste the martingale measures together. This equivalence, however, breaks down in the presence of transaction cost. A simple example is the two-period market: $\underline{S}_0 = 1, \bar{S}_0 = 3, \underline{S}_1 = 2, \bar{S}_1 = 4, \underline{S}_2 = 3.5, \bar{S}_2 = 5$ where $\underline{S}_t, \bar{S}_t$ are the bid and ask prices, respectively. Each period is arbitrage-free, but buying at time 0 and selling at time 2 is an arbitrage for the two-period market. So we cannot in general paste two one-period martingale measures to get a two-period martingale measure; in particular, the endpoints of the underlying martingales constructed for each single period may not match. We need to solve a non-dominated martingale selection problem. The martingale selection problem when \mathcal{P} is a singleton was studied by *Rokhlin* in a series of papers [65, 67, 66] using the notion of support of regular conditional upper distribution of set-valued maps. In our case, it is difficult to talk about conditional distribution due to the lack of dominating measure. Nevertheless, we got some inspiration from [67, 66] and [75] and developed a backward-forward scheme:

Backward recursion

Modify the original bid-ask prices backward in time by potentially more favorable ones to account for the missing future investment opportunities.

Forward extension

Extend the dual element forward in time in the modified market.

Unfortunately, when there is no dominating measure, the backward recursion brings some measurability issues. We overcame these issues by making a suitable

continuity assumption. The necessity of this assumption is briefly discussed in Section 2.3.3. Our contribution can be seen as a particular extension (in the two-dimensional case and under the additional continuity assumption) of the FTAP in [25] to the frictional case, as well as a generalization of [66] on the martingale selection problem to the non-dominated case. We also give an existence result of the optimal super-hedging strategy when $\text{NA}^r(\mathcal{P})$ holds. The challenges of a super-hedging duality and multi-asset extension is discussed in Section 2.4.

We end the introduction with a mathematical description of model uncertainty which will be used throughout this chapter. The notations are taken from [25].

2.1.1 The uncertainty set \mathcal{P}

Let $T \in \mathbb{N}$ be the time horizon and let Ω_1 be a Polish space. Let $\Omega_t := \Omega_1^t$ be the t -fold Cartesian product with the convention that Ω_0 is a singleton. Denote by $\mathcal{B}(\Omega_t)$ the Borel sigma-algebra on Ω_t , and by \mathcal{F}_t the universal completion of $\mathcal{B}(\Omega_t)$. We write (Ω, \mathcal{F}) for $(\Omega_T, \mathcal{F}_T)$. Let $\mathfrak{P}(\Omega_1)$ denote the set of all probability measures on $(\Omega_1, \mathcal{B}(\Omega_1))$. For each $t \in \{0, \dots, T-1\}$ and $\omega \in \Omega_t$, we are given a nonempty convex set $\mathcal{P}_t(\omega) \subseteq \mathfrak{P}(\Omega_1)$, representing the set of possible models for the $(t+1)$ -th period. We assume the graph of \mathcal{P}_t (considered as a set valued map from Ω_t to $\mathfrak{P}(\Omega_1)$) is analytic. This assumption ensures that \mathcal{P}_t admits a universally measurable selector, which we will denote by P_t . Define the uncertainty set $\mathcal{P} \subseteq \mathfrak{P}(\Omega)$ of the multi-period market by

$$\mathcal{P} := \{P_0 \otimes \dots \otimes P_{T-1} : \text{each } P_t \text{ is a universally measurable selector of } \mathcal{P}_t\},$$

where for $A \in \mathcal{F}$,

$$P_0 \otimes \dots \otimes P_{T-1}(A) = \int_{\Omega_1} \dots \int_{\Omega_1} 1_A(\omega_1, \dots, \omega_T) P_{T-1}(\omega_1, \dots, \omega_{T-1}; d\omega_T) \dots P_0(d\omega_1).$$

2.2 When bid-ask spread is on hedging options only

2.2.1 Fundamental theorem with robust no-arbitrage

Let $S_t = (S_t^1, \dots, S_t^d)$ be the prices of d traded stocks at time $t \in \{0, 1, \dots, T\}$ and \mathcal{H} be the set of all predictable \mathbb{R}^d -valued processes, which will serve as our trading strategies. Let $g = (g^1, \dots, g^e)$ be the payoff of e options that mature at time T , and can be traded only at time zero with bid price \underline{g} and ask price \bar{g} , with $\bar{g} \geq \underline{g}$ (the inequality holds component-wise). g^1, \dots, g^e will be referred to as hedging options. We assume S_t and g are Borel measurable, and there are no transaction costs in the trading of stocks. We also assume the risk-free rate is zero.

Definition 2.2.1 (No-arbitrage and robust no-arbitrage). We say that condition $\text{NA}(\mathcal{P})$ holds if for all $(H, h) \in \mathcal{H} \times \mathbb{R}^e$,

$$H \bullet S_T + h^+(g - \bar{g}) - h^-(g - \underline{g}) \geq 0 \quad \mathcal{P} - \text{quasi-surely (-q.s.)}^1$$

implies

$$H \bullet S_T + h^+(g - \bar{g}) - h^-(g - \underline{g}) = 0 \quad \mathcal{P}\text{-q.s.}$$

where $H \bullet S_t = \sum_{u=1}^t H_u(S_u - S_{u-1})$ is the discrete-time integral, and h^\pm are the usual (component-wise) positive/negative part of h .²

We say that condition $\text{NA}^r(\mathcal{P})$ holds if there exists \underline{g}', \bar{g}' such that $[\underline{g}', \bar{g}'] \subseteq \text{ri}[\underline{g}, \bar{g}]$ and $\text{NA}(\mathcal{P})$ holds if g has bid-ask prices \underline{g}', \bar{g}' .³

Definition 2.2.2 (Super-hedging price). For a given a random variable f , its super-

¹A set is \mathcal{P} -polar if it is P -null for all $P \in \mathcal{P}$. A property is said to hold \mathcal{P} -q.s. if it holds outside a \mathcal{P} -polar set.

²When we multiply two vectors, we mean their inner product.

³“ri” stands for relative interior. $[\underline{g}', \bar{g}'] \subseteq \text{ri}[\underline{g}, \bar{g}]$ means component-wise inclusion.

hedging price is defined as

$$\begin{aligned} \pi(f) &:= \inf\{x \in \mathbb{R} : \exists (H, h) \in \mathcal{H} \times \mathbb{R}^e \text{ such that} \\ &\quad x + H \bullet S_T + h^+(g - \bar{g}) - h^-(g - \underline{g}) \geq f \text{ } \mathcal{P}\text{-q.s.}\} \end{aligned}$$

Any pair $(H, h) \in \mathcal{H} \times \mathbb{R}^e$ in the above definition is called a semi-static hedging strategy.

Remark 2.2.3. [1] Let $\hat{\pi}(g^i)$ and $\hat{\pi}(-g^i)$ be the super-hedging prices of g^i and $-g^i$, where the hedging is done using stocks and options excluding g^i . $\text{NA}^r(\mathcal{P})$ implies either

$$-\hat{\pi}(-g^i) \leq \underline{g}^i = \bar{g}^i \leq \hat{\pi}(g^i)$$

or

$$(2.1) \quad -\hat{\pi}(-g^i) \leq (\bar{g}')^i < \bar{g}^i \quad \text{and} \quad \underline{g}^i < (\underline{g}')^i \leq \hat{\pi}(g^i)$$

where \underline{g}', \bar{g}' are the more favorable bid-ask prices in the definition of robust no-arbitrage. The reason for working with robust no-arbitrage is to be able to have the strictly inequalities in (2.1) for options with non-zero spread, which turns out to be crucial in the proof of the closedness of the set of hedgeable claims in (2.3) (hence the existence of an optimal hedging strategy), as well as in the construction of a dual element (see (2.6)).

[2] Clearly $\text{NA}^r(\mathcal{P})$ implies $\text{NA}(\mathcal{P})$, but the converse is not true. For example, assume in the market there is no stock, and there are only two options: $g_1(\omega) = g_2(\omega) = \omega$, $\omega \in \Omega := [0, 1]$. Let \mathcal{P} be the set of probability measures on Ω , $\underline{g}_1 = \bar{g}_1 = 1/2$, $\underline{g}_2 = 1/4$ and $\bar{g}_2 = 1/2$. Then $\text{NA}(\mathcal{P})$ holds while $\text{NA}^r(\mathcal{P})$ fails.

For $b, a \in \mathbb{R}^e$, let

$$\mathcal{Q}^{[b, a]} := \{Q \lll \mathcal{P} : Q \text{ is a martingale measure and } E^Q[g] \in [b, a]\}$$

where $Q \lll P$ means $\exists P \in \mathcal{P}$ such that $Q \ll P$.⁴ Let $\mathcal{Q}_\varphi^{[b,a]} := \{Q \in \mathcal{Q} : E^Q[\varphi] < \infty\}$. When $[b, a] = [\underline{g}, \bar{g}]$, we drop the superscript and simply write $\mathcal{Q}, \mathcal{Q}_\varphi$. Also define

$$\mathcal{Q}^s := \{Q \lll P : Q \text{ is a martingale measure and } E^Q[g] \in ri[\underline{g}, \bar{g}]\}$$

and $\mathcal{Q}_\varphi^s := \{Q \in \mathcal{Q}^s : E^Q[\varphi] < \infty\}$.

Theorem 2.2.4. *Let $\varphi \geq 1$ be a random variable such that $|g^i| \leq \varphi \forall i = 1, \dots, e$. The following statements hold:*

(a) (FTAP): *The following statements are equivalent*

(i) *$NA^r(\mathcal{P})$ holds.*

(ii) *There exists $[\underline{g}', \bar{g}'] \subseteq ri[\underline{g}, \bar{g}]$ such that $\forall P \in \mathcal{P}, \exists Q \in \mathcal{Q}_\varphi^{[\underline{g}', \bar{g}]}$ such that $P \ll Q$.*

(b) (Super-hedging) *Suppose $NA^r(\mathcal{P})$ holds. Let $f : \Omega \rightarrow \mathbb{R}$ be Borel measurable such that $|f| \leq \varphi$. The super-hedging price is given by*

$$(2.2) \quad \pi(f) = \sup_{Q \in \mathcal{Q}_\varphi^s} E^Q[f] = \sup_{Q \in \mathcal{Q}_\varphi} E^Q[f] \in (-\infty, \infty]$$

and there exists $(H, h) \in \mathcal{H} \times \mathbb{R}^e$ such that $\pi(f) + H \bullet S_T + h^+(g - \bar{g}) - h^-(g - \underline{g}) \geq f$ \mathcal{P} -q.s..

Proof. It is easy to show (ii) in (a) implies that $NA(\mathcal{P})$ holds for the market with bid-ask prices \underline{g}', \bar{g}' , Hence $NA^r(\mathcal{P})$ holds for the original market. The rest of our proof consists two parts as follows.

Part 1: $\pi(f) > -\infty$ and the existence of an optimal hedging strategy in

⁴ $E^Q[g] \in [b, a]$ means $E^Q[g^i] \in [b^i, a^i]$ for all $i = 1, \dots, e$.

(b). Once we show that the set

$$(2.3) \quad \mathcal{C}_g := \{H \bullet S_T + h^+(g - \bar{g}) - h^-(g - \underline{g}) : (H, h) \in \mathcal{H} \times \mathbb{R}^e\} - \mathcal{L}_+^0$$

is \mathcal{P} -*q.s.* closed (*i.e.*, if $(W^n)_{n=1}^\infty \subset \mathcal{C}_g$ and $W^n \rightarrow W$ \mathcal{P} -*q.s.*, then $W \in \mathcal{C}_g$), the argument used in the proof of ([25], [Theorem 2.3]) would conclude the results in part 1. We will demonstrate the closedness of \mathcal{C}_g in the rest of this part.

Write $g = (u, v)$, where $u = (g^1, \dots, g^r)$ consists of the hedging options without bid-ask spread, *i.e.*, $\underline{g}^i = \bar{g}^i$ for $i = 1, \dots, r$, and $v = (g^{r+1}, \dots, g^e)$ consists of those with spread, *i.e.*, $\underline{g}^i < \bar{g}^i$ for $i = r+1, \dots, e$, for some $r \in \{0, \dots, e\}$. Denote $\underline{u} := (\underline{g}^1, \dots, \underline{g}^r)$ and similarly for \underline{v} and \bar{v} . Define

$$\mathcal{C} := \{H \bullet S_T + \alpha(u - \underline{u}) : (H, \alpha) \in \mathcal{H} \times \mathbb{R}^r\} - \mathcal{L}_+^0$$

Then \mathcal{C} is \mathcal{P} -*q.s.* closed by ([25], [Theorem 2.2]).

Let $W^n \rightarrow W$ \mathcal{P} -*q.s.* with

$$(2.4) \quad W^n = H^n \bullet S_T + \alpha^n(u - \underline{u}) + (\beta^n)^+(v - \bar{v}) - (\beta^n)^-(v - \underline{v}) - U^n \in \mathcal{C}_g$$

where $(H^n, \alpha^n, \beta^n) \in \mathcal{H} \times \mathbb{R}^r \times \mathbb{R}^{e-r}$ and $U^n \in \mathcal{L}_+^0$. If $(\beta^n)_n$ is not bounded, then by passing to subsequence if necessary, we may assume that $0 < \|\beta^n\| \rightarrow \infty$ and rewrite (2.4) as

$$\frac{H^n}{\beta^n} \bullet S_T + \frac{\alpha^n}{\|\beta^n\|}(u - \underline{u}) \geq \frac{W^n}{\|\beta^n\|} - \left(\frac{\beta^n}{\|\beta^n\|}\right)^+(v - \bar{v}) + \left(\frac{\beta^n}{\|\beta^n\|}\right)^-(v - \underline{v}) \in \mathcal{C}$$

where $\|\cdot\|$ represents the sup-norm. Since \mathcal{C} is \mathcal{P} -*q.s.* closed, the limit of the right hand side above is also in \mathcal{C} , *i.e.*, there exists some $(H, \alpha) \in \mathcal{H} \times \mathbb{R}^r$, such that

$$H \bullet S_T + \alpha(u - \underline{u}) \geq -\beta^+(v - \bar{v}) + \beta^-(v - \underline{v}), \quad \mathcal{P} - a.s.$$

where β is the limit of $(\beta^n)_n$ along some subsequence with $\|\beta\| = 1$. $\text{NA}(\mathcal{P})$ implies that

$$(2.5) \quad H \bullet S_T + \alpha(u - \underline{u}) + \beta^+(v - \bar{v}) - \beta^-(v - \underline{v}) = 0, \quad \mathcal{P} - a.s.$$

As $\beta =: (\beta_{r+1}, \dots, \beta_e) \neq 0$, we assume without loss of generality (w.l.o.g.) that $\beta_e \neq 0$. If $\beta_e < 0$, then we have from (2.5) that

$$\underline{g}^e + \frac{H}{\beta_e^-} \bullet S_T + \frac{\alpha}{\beta_e^-} (u - \underline{u}) + \sum_{i=r+1}^{e-1} \left[\frac{\beta_i^+}{\beta_e^-} (g^i - \bar{g}^i) - \frac{\beta_i^-}{\beta_e^-} (g^i - \underline{g}^i) \right] = g^e, \quad \mathcal{P} - a.s.$$

Therefore $\hat{\pi}(g^e) \leq \underline{g}_e$, which contradicts the robust no-arbitrage property (see (2.1)) of g^e . Here $\hat{\pi}(g^e)$ is the super-hedging price of g^e using S and g excluding g^e . Similarly we get a contradiction if $\beta_e > 0$.

Thus $(\beta^n)_n$ is bounded, and has a limit $\beta \in \mathbb{R}^{e-r}$ along some subsequence $(n_k)_k$. Since by (2.4)

$$H^n \bullet S_T + \alpha^n (u - \underline{u}) \geq W^n - (\beta^n)^+(v - \bar{v}) + (\beta^n)^-(v - \underline{v}) \in \mathcal{C}$$

the limit of the right hand side above along $(n_k)_k$, $W - \beta^+(v - \bar{v}) + \beta^-(v - \underline{v})$, is also in \mathcal{C} by its closedness, which implies $W \in \mathcal{C}_g$.

Part 2: (i) \Rightarrow (ii) in part (a) and (2.13) in part (b). We will prove the results by an induction on the number of hedging options, as in ([25], [Theorem 5.1]). Suppose the results hold for the market with options g^1, \dots, g^e . We now introduce an additional option $f \equiv g^{e+1}$ with $|f| \leq \varphi$, available at bid-ask prices $\underline{f} < \bar{f}$ at time zero. (When the bid and ask prices are the same for f , then the proof is identical to [25].)

(i) \implies (ii) in (a): Let $\pi(f)$ be the super-hedging price when stocks and g^1, \dots, g^e are available for trading. By $\text{NA}^r(\mathcal{P})$ and (2.13) in part (b) of the induction hypothesis, we have

$$(2.6) \quad \bar{f} > \bar{f}' \geq -\pi(-f) = \inf_{Q \in \mathcal{Q}_\varphi^s} E^Q[f] \quad \text{and} \quad \underline{f} < \underline{f}' \leq \pi(f) = \sup_{Q \in \mathcal{Q}_\varphi^s} E^Q[f]$$

where $[\underline{f}', \bar{f}'] \subseteq (\underline{f}, \bar{f})$ comes from the definition of robust no-arbitrage. This implies that there exists $Q_+, Q_- \in \mathcal{Q}_\varphi^s$ such that $E^{Q_+}[f] > \underline{f}''$ and $E^{Q_-}[f] < \bar{f}''$ where $\underline{f}'' =$

$\frac{1}{2}(\underline{f} + \underline{f}'), \bar{f}'' = \frac{1}{2}(\bar{f} + \bar{f}')$. By (a) of induction hypothesis, there exists $[b, a] \subseteq ri[\underline{g}, \bar{g}]$ such that for any $P \in \mathcal{P}$, we can find $Q_0 \in \mathcal{Q}_\varphi^{[b,a]}$ satisfying $P \ll Q_0 \lll \mathcal{P}$. Define

$$\underline{g}' = \min(b, E^{Q_+}[g], E^{Q_-}[g]), \quad \text{and} \quad \bar{g}' = \max(a, E^{Q_+}[g], E^{Q_-}[g])$$

where the minimum and maximum are taken component-wise. We have $[b, a] \subseteq [\underline{g}', \bar{g}'] \subseteq ri[\underline{g}, \bar{g}]$ and $Q_+, Q_- \in \mathcal{Q}_\varphi^{[\underline{g}', \bar{g}]}$.

Now, let $P \in \mathcal{P}$. (a) of induction hypothesis implies the existence of a $Q_0 \in \mathcal{Q}_\varphi^{[b,a]} \subseteq \mathcal{Q}_\varphi^{[\underline{g}', \bar{g}]}$ satisfying $P \ll Q_0 \lll \mathcal{P}$. Define

$$Q := \lambda_- Q_- + \lambda_0 Q_0 + \lambda_+ Q_+$$

Then $Q \in \mathcal{Q}_\varphi^{[\underline{g}', \bar{g}]}$ and $P \ll Q$. By choosing suitable weights $\lambda_-, \lambda_0, \lambda_+ \in (0, 1), \lambda_- + \lambda_0 + \lambda_+ = 1$, we can make $E^Q[f] \in [\underline{f}'', \bar{f}''] \subseteq ri[\underline{f}, \bar{f}]$.

(2.13) in (b): Let ξ be a Borel measurable function such that $|\xi| \leq \varphi$. Write $\pi'(\xi)$ for its super-hedging price when stocks and $g^1, \dots, g^e, f \equiv g^{e+1}$ are traded, $\mathcal{Q}'_\varphi := \{Q \in \mathcal{Q}_\varphi : E^Q[f] \in [\underline{f}, \bar{f}]\}$ and $\mathcal{Q}^s_\varphi := \{Q \in \mathcal{Q}_\varphi^s : E^Q[f] \in (\underline{f}, \bar{f})\}$. We want to show

$$(2.7) \quad \pi'(\xi) = \sup_{Q \in \mathcal{Q}^s_\varphi} E^Q[\xi] = \sup_{Q \in \mathcal{Q}'_\varphi} E^Q[\xi]$$

It is easy to see that

$$(2.8) \quad \pi'(\xi) \geq \sup_{Q \in \mathcal{Q}'_\varphi} E^Q[\xi] \geq \sup_{Q \in \mathcal{Q}^s_\varphi} E^Q[\xi]$$

and we shall focus on the reverse inequalities. Let us assume first that ξ is bounded from above, and thus $\pi'(\xi) < \infty$. By a translation we may assume $\pi'(\xi) = 0$.

First, we show $\pi'(\xi) \leq \sup_{Q \in \mathcal{Q}'_\varphi} E^Q[\xi]$. It suffices to show the existence of a sequence $\{Q_n\} \subseteq \mathcal{Q}'_\varphi$ such that $\lim_n E^{Q_n}[f] \in [\underline{f}, \bar{f}]$ and $\lim_n E^{Q_n}[\xi] = \pi'(\xi) = 0$. (See page 30 of [25] for why this is sufficient.) In other words, we want to show that

$$\overline{\{E^Q[(f, \xi)] : Q \in \mathcal{Q}'_\varphi\}} \cap ([\underline{f}, \bar{f}] \times \{0\}) \neq \emptyset$$

Suppose the above intersection is empty. Then there exists a vector $(y, z) \in \mathbb{R}^2$ with $|(y, z)| = 1$ that strictly separates the two closed, convex sets, *i.e.*, there exists $b \in \mathbb{R}$ s.t.

$$(2.9) \quad \sup_{Q \in \mathcal{Q}_\varphi} E^Q[yf + z\xi] < b \quad \text{and} \quad \inf_{a \in [\underline{f}, \bar{f}]} ya > b$$

It follows that

$$(2.10) \quad y^+ \underline{f} - y^- \bar{f} + \pi'(z\xi) \leq \pi'(yf + z\xi) \leq \pi(yf + z\xi) = \sup_{Q \in \mathcal{Q}_\varphi} E^Q[yf + z\xi] < b < y^+ \underline{f} - y^- \bar{f},$$

where the first inequality is because one can super-replicate $z\xi = (yf + z\xi) + (-yf)$ from initial capital $\pi'(yf + z\xi) - y^+ \underline{f} + y^- \bar{f}$, the second inequality is due to the fact that having more options to hedge reduces hedging cost, and the middle equality is by (b) of induction hypothesis. The last two inequalities are due to (2.9).

It follows from (2.10) that $\pi'(z\xi) < 0$. Therefore, we must have that $z < 0$, otherwise $\pi'(z\xi) = z\pi'(\xi) = 0$ (since the super-hedging price is positively homogenous). Recall that we have proved in part (a) that $\mathcal{Q}'_\varphi \neq \emptyset$. Let $Q' \in \mathcal{Q}'_\varphi \subseteq \mathcal{Q}_\varphi$. The part of (2.10) after the equality implies that $yE^{Q'}[f] + zE^{Q'}[\xi] < y^+ \underline{f} - y^- \bar{f}$. Since $E^{Q'}[f] \in [\underline{f}, \bar{f}]$, we get $zE^{Q'}[\xi] < y^+(\underline{f} - E^{Q'}[f]) - y^-(\bar{f} - E^{Q'}[f]) \leq 0$. Since $z < 0$, $E^{Q'}[\xi] > 0$. But by (2.8), $E^{Q'}[\xi] \leq \pi'(\xi) = 0$, which is a contradiction.

Next, we show $\sup_{Q \in \mathcal{Q}'_\varphi} E^Q[\xi] \leq \sup_{Q \in \mathcal{Q}'^s_\varphi} E^Q[\xi]$. It suffices to show for any $\varepsilon > 0$ and every $Q \in \mathcal{Q}'_\varphi$, we can find $Q^s \in \mathcal{Q}'^s_\varphi$ such that $E^{Q^s}[\xi] > E^Q[\xi] - \varepsilon$. To this end, let $Q' \in \mathcal{Q}'_\varphi$ which is nonempty by part (a). Define

$$Q^s := (1 - \lambda)Q + \lambda Q'$$

We have $Q^s \lll \mathcal{P}$ by the convexity of \mathcal{P} , and $Q^s \in \mathcal{Q}'^s_\varphi$ if $\lambda \in (0, 1]$. Moreover,

$$E^{Q^s}[\xi] = (1 - \lambda)E^Q[\xi] + \lambda E^{Q'}[\xi] \rightarrow E^Q[\xi] \quad \text{as } \lambda \rightarrow 0$$

So for $\lambda > 0$ sufficiently close to zero, the Q^s constructed above satisfies $E^{Q^s}[\xi] > E^Q[\xi] - \varepsilon$. Hence, we have shown that the supremum over \mathcal{Q}'_φ and \mathcal{Q}^s_φ are equal. This finishes the proof for upper bounded ξ .

Finally, when ξ is not bounded from above, we can apply the previous result to $\xi \wedge n$, and then let $n \rightarrow \infty$ and use the closedness of \mathcal{C}_g in (2.3) to show that (2.13) holds. The argument would be the same as the last paragraph in the proof of [25, Theorem 3.4] and we omit it here. \square

2.2.2 A sharper fundamental theorem with a non-redundancy assumption

We now introduce the concept of non-redundancy. With this additional assumption we will sharpen our result.

Definition 2.2.5 (Non-redundancy). A hedging option g^i is said to be non-redundant if it is not perfectly replicable by stocks and other hedging options, *i.e.*, there does not exist $x \in \mathbb{R}$ and a semi-static hedging strategy $(H, h) \in \mathcal{H} \times \mathbb{R}^e$ such that

$$x + H \bullet S_T + \sum_{j \neq i} h^j g^j = g^i \text{ } \mathcal{P}\text{-q.s.}$$

Remark 2.2.6. $\text{NA}^r(\mathcal{P})$ does not imply non-redundancy. For Instance, having only two identical options in the market whose payoffs are in $[c, d]$, with identical bid-ask prices b and a satisfying $b < c$ and $a > d$, would give a trivial counter example where $\text{NA}^r(\mathcal{P})$ holds yet we have redundancy.

Lemma 2.2.7. *Suppose all hedging options with non-zero spread are non-redundant. Then $\text{NA}(\mathcal{P})$ implies $\text{NA}^r(\mathcal{P})$.*

Proof. Let $g = (g^1, \dots, g^{r+s})$, where $u := (g^1, \dots, g^r)$ consists of the hedging options without bid-ask spread, *i.e.*, $\underline{g}^i = \bar{g}^i$ for $i = 1, \dots, r$, and $(g^{r+1}, \dots, g^{r+s})$ consists of

those with bid-ask spread, *i.e.*, $\underline{g}^i < \bar{g}^i$ for $i = r + 1, \dots, r + s$. We shall prove the result by induction on s . Obviously the result holds when $s = 0$. Suppose the result holds for $s = k \geq 0$. Then for $s = k + 1$, denote $v := (g^{r+1}, \dots, g^{r+k})$, $\underline{v} := (\underline{g}^{r+1}, \dots, \underline{g}^{r+k})$ and $\bar{v} := (\bar{g}^{r+1}, \dots, \bar{g}^{r+k})$. Denote $f := g^{r+k+1}$.

By the induction hypothesis, there exists $[\underline{v}', \bar{v}'] \subset (\underline{v}, \bar{v})$ be such that $\text{NA}(\mathcal{P})$ holds in the market with stocks, options u and options v with any bid-ask prices b and a satisfying $[\underline{v}', \bar{v}'] \subset [b, a] \subset (\underline{v}, \bar{v})$. Let $\underline{v}_n \in (\underline{v}, \underline{v}')$, $\bar{v}_n \in (\bar{v}', \bar{v})$, $\underline{f}_n > \underline{f}$ and $\bar{f}_n < \bar{f}$, such that $\underline{v}_n \searrow \underline{v}$, $\bar{v}_n \nearrow \bar{v}$, $\underline{f}_n \searrow \underline{f}$ and $\bar{f}_n \nearrow \bar{f}$. We shall show that for some n , $\text{NA}(\mathcal{P})$ holds with stocks, options u , options v with bid-ask prices \underline{v}_n and \bar{v}_n , option f with bid-ask prices \underline{f}_n and \bar{f}_n . We will show it by contradiction.

If not, then for each n , there exists $(H^n, h_u^n, h_v^n, h_f^n) \in \mathcal{H} \times \mathbb{R}^r \times \mathbb{R}^k \times \mathbb{R}$ such that

$$(2.11)$$

$$H^n \bullet S_T + h_u^n (u - \underline{u}) + (h_v^n)^+ (v - \bar{v}_n) - (h_v^n)^- (v - \underline{v}_n) + (h_f^n)^+ (f - \bar{f}_n) - (h_f^n)^- (f - \underline{f}_n) \geq 0, \mathcal{P} - q.s.$$

and the strict inequality for the above holds with positive probability under some $P_n \in \mathcal{P}$. Hence $h_f^n \neq 0$. By a normalization, we can assume that $|h_f^n| = 1$. By extracting a subsequence, we can w.l.o.g. assume that $h_f^n = -1$ (the argument when assuming $h_f^n = 1$ is similar). If $(h_u^n, h_v^n)_n$ is not bounded, then w.l.o.g. we assume that $0 < c^n := \|(h_u^n, h_v^n)\| \rightarrow \infty$. By (2.11) we have that

$$\frac{H^n}{c^n} \bullet S_T + \frac{h_u^n}{c^n} (u - \underline{u}) + \frac{(h_v^n)^+}{c^n} (v - \bar{v}_n) - \frac{(h_v^n)^-}{c^n} (v - \underline{v}_n) - \frac{1}{c^n} (f - \underline{f}_n) \geq 0, \mathcal{P} - q.s.$$

By ([25], [Theorem 2.2]), there exists $H \in \mathcal{H}$, such that

$$H \bullet S_T + h_u (u - \underline{u}) + h_v^+ (v - \bar{v}) - h_v^- (v - \underline{v}) \geq 0, \mathcal{P} - q.s.$$

where (h_u, h_v) is the limit of $(h_u^n/c^n, h_v^n/c^n)$ along some subsequence with $\|(h_u, h_v)\| =$

1. $\text{NA}(\mathcal{P})$ implies that

$$(2.12) \quad H \bullet S_T + h_u (u - \underline{u}) + h_v^+ (v - \bar{v}) - h_v^- (v - \underline{v}) = 0, \mathcal{P} - q.s.$$

Since $(h_u, h_v) \neq 0$, (2.12) contradicts the non-redundancy assumption of (u, v) .

Therefore, $(h_u^n, h_v^n)_n$ is bounded, and w.l.o.g. assume it has the limit (\hat{h}_u, \hat{h}_v) .

Then applying ([25], [Theorem 2.2]) in (2.11), there exists $\hat{H} \in \mathcal{H}$ such that

$$\hat{H} \bullet S_T + \hat{h}_u(u - \underline{u}) + \hat{h}_v^+(v - \bar{v}) - \hat{h}_v^-(v - \underline{v}) - (f - \underline{f}) \geq 0, \mathcal{P} - q.s.$$

$\text{NA}(\mathcal{P})$ implies that

$$\hat{H} \bullet S_T + \hat{h}_u(u - \underline{u}) + \hat{h}_v^+(v - \bar{v}) - \hat{h}_v^-(v - \underline{v}) - (f - \underline{f}) = 0, \mathcal{P} - q.s.$$

which contradicts the non-redundancy assumption of f . \square

We have the following FTAP and super-hedging result in terms of $\text{NA}(\mathcal{P})$ instead of $\text{NA}^r(\mathcal{P})$, when we additionally assume the non-redundancy of g .

Theorem 2.2.8. *Suppose all hedging options with non-zero spread are non-redundant. Let $\varphi \geq 1$ be a random variable such that $|g^i| \leq \varphi \forall i = 1, \dots, e$. The following statements hold:*

(a') (FTAP): *The following statements are equivalent*

(i) *$\text{NA}(\mathcal{P})$ holds.*

(ii) *$\forall P \in \mathcal{P}, \exists Q \in \mathcal{Q}_\varphi$ such that $P \ll Q$.*

(b') (Super-hedging) *Suppose $\text{NA}(\mathcal{P})$ holds. Let $f : \Omega \rightarrow \mathbb{R}$ be Borel measurable such that $|f| \leq \varphi$. The super-hedging price is given by*

$$(2.13) \quad \pi(f) = \sup_{Q \in \mathcal{Q}_\varphi} E^Q[f] \in (-\infty, \infty]$$

and there exists $(H, h) \in \mathcal{H} \times \mathbb{R}^e$ such that $\pi(f) + H \bullet S_T + h^+(g - \bar{g}) - h^-(g - \underline{g}) \geq f$ \mathcal{P} -q.s..

Proof. (a')(ii) \implies (a')(i) is trivial. Now if (a')(i) holds, then by Lemma 2.2.7, (a)(i) in Theorem 2.2.4 holds, which implies (a)(ii) holds, and thus (a')(ii) holds. Finally, (b') is implied by Lemma 2.2.7 and Theorem 2.2.4(b). \square

Remark 2.2.9. Theorem 2.2.8 generalizes the results of [25] to the case when the option prices are quoted with bid-ask spreads. When \mathcal{P} is the set of all probability measures and the given options are all call options written on the dynamically traded assets, a result with option bid-ask spreads similar to Theorem 2.2.8-(a) had been obtained by [28]; see Proposition 4.1 therein, although the non-redundancy condition did not actually appear. (The objective of [28] was to obtain relationships between the option prices which are necessary and sufficient to rule out semi-static arbitrage and the proof relies on determining the correct set of relationships and then identifying a martingale measure.)

However, the no arbitrage concept used in [28] is different: the author there assumes that there is no *weak arbitrage* in the sense of [31]; see also [33] and [1].⁵ (Recall that a market is said to have weak arbitrage if for any given model (probability measure) there is an arbitrage strategy which is an arbitrage in the classical sense.) The arbitrage concept used here and in [25] is weaker, in that we say that a non-negative wealth (\mathcal{P} -q.s.) is an arbitrage even if there is a single P under which the wealth process is a classical arbitrage. Hence our no-arbitrage condition is stronger than the one used in [28]. But what we get out from a stronger assumption is the existence of a martingale measure $Q \in \mathcal{Q}_\varphi$ for each $P \in \mathcal{P}$. Whereas [28] only guarantees the existence of only one martingale measure which prices the hedging options correctly.

⁵The no-arbitrage assumption in [1] is the model independent arbitrage of [31]. However that paper rules out the model dependent arbitrage by assuming that a superlinearly growing option can be bought for static hedging.

2.3 When bid-ask spread is on one dynamically traded asset

2.3.1 The market model and main results

Consider a financial market consisting of a money market account with zero interest rate, and a stock with bid price \underline{S}_t and ask price \bar{S}_t .⁶ $\underline{S}_t, \bar{S}_t : \Omega_t \rightarrow \mathbb{R}$ are assumed to be positive, universally measurable for all t . Later on, we will replace universal measurability by continuity. We shall often use square and curly brackets to denote the stock market: $[\underline{S}, \bar{S}]$ refers to a multi-period market with bid price \underline{S}_t and ask price \bar{S}_t for all t and for any $\omega \in \Omega_t$, $\{[\underline{S}_t(\omega), \bar{S}_t(\omega)]; [\underline{S}_{t+1}(\omega, \cdot), \bar{S}_{t+1}(\omega, \cdot)]\}$ refers to the one-period market with bid price $\underline{S}_t(\omega)$ and ask price $\bar{S}_t(\omega)$ at time t , and bid price $\underline{S}_{t+1}(\omega, \cdot)$ and ask price $\bar{S}_{t+1}(\omega, \cdot)$ at time $t + 1$. When $\underline{S}_t(\omega) = \bar{S}_t(\omega) = s$, we simply write $\{s; [\underline{S}_{t+1}(\omega, \cdot), \bar{S}_{t+1}(\omega, \cdot)]\}$ for the one-period market. The solvency cone K_t at time t is the closed convex cone in \mathbb{R}^2 spanned by the unit vectors e_1, e_2 and the vectors $\bar{S}_t e_1 - e_2, \frac{1}{\underline{S}_t} e_2 - e_1$. That is, K_t is the cone of portfolios that can be liquidated into the zero portfolio; it contains the nonnegative orthant $\mathbb{R}_{\geq 0}^2$. $-K_t$ is the cone of portfolios available at price zero.

A *self-financing*⁷ portfolio process is an \mathbb{R}^2 -valued predictable process $\phi = (\phi^0, \phi^1)$ satisfying $\Delta\phi_{t+1} := \phi_{t+1} - \phi_t \in -K_t$ for all $t = 0, \dots, T$. We always define $\phi_0 = 0$. An equivalent expression for $\Delta\phi_{t+1} \in -K_t$ is $\Delta\phi_{t+1}^0 \leq -(\Delta\phi_{t+1}^1)^+ \bar{S}_t + (\Delta\phi_{t+1}^1)^- \underline{S}_t$. Denote by \mathcal{H} the set of self-financing portfolio processes. Let A_T be the set of all \mathbb{R}^2 -valued, \mathcal{F}_T -measurable functions which \mathcal{P} -q.s. equal to ϕ_{T+1} for some $\phi \in \mathcal{H}$. A_T is interpreted as the set of hedgeable claims (in terms of physical units) from zero initial endowment. It is easy to see that A_T is solid in the sense that if $f \in A_T$ and

⁶Generally, if there are two dynamically traded assets having exchange rates $\pi^{12}\pi^{21} \geq 1$ where π^{ij} is the number of physical units of the i -th asset needed to exchange for one unit of the j -th asset, then one can treat asset 1 as the numéraire, and define the bid and ask prices of asset 2 in terms of asset 1 as $\underline{S} = 1/\pi^{21}$ and $\bar{S} = \pi^{12}$, respectively.

⁷We allow agents to throw away non-negative quantities of the assets.

$g \leq f$ \mathcal{P} -q.s. (the inequality holds componentwise), then $g \in A_T$.

Definition 2.3.1.

- (i) Condition $\text{NA}(\mathcal{P})$ holds if for all $f \in A_T$, $f \geq 0$ \mathcal{P} -q.s. implies $f = 0$ \mathcal{P} -q.s..
- (ii) Condition $\text{NA}^r(\mathcal{P})$ holds if there exist bid-ask price processes $\underline{S}'_t, \overline{S}'_t$ with the same measurability and continuity property as $\underline{S}_t, \overline{S}_t$ such that $[\underline{S}'_t, \overline{S}'_t] \subseteq \text{ri}[\underline{S}_t, \overline{S}_t]$ and $\text{NA}(\mathcal{P})$ holds for the $[\underline{S}', \overline{S}']$ -market.⁸

It is clear that $\text{NA}^r(\mathcal{P})$ implies $\text{NA}(\mathcal{P})$ since any hedgeable claim is also hedgeable in markets with smaller friction, and they are equivalent when the market is frictionless on a \mathcal{P} -q.s. set.

Remark 2.3.2. $\text{NA}(P) \forall P \in \mathcal{P}$ implies $\text{NA}(\mathcal{P})$. Indeed, let $f \in A_T$ be such that $f \geq 0$ \mathcal{P} -q.s. hence P -a.s. for all $P \in \mathcal{P}$. For each P , $\text{NA}(P)$ implies $f = 0$ P -a.s.. Since this holds for all $P \in \mathcal{P}$, $f = 0$ \mathcal{P} -q.s. and $\text{NA}(\mathcal{P})$ holds. The reverse direction is not true. Consider a one-period market with $S_0 = 2$ and $S_1(\omega) = 1, S_1(\omega') = 3$. Let $P_1 = \delta_\omega, P_2 = \delta_{\omega'}$ be the Dirac measures concentrated on ω and ω' , respectively. Then it is easy to see that there is arbitrage under both P_1 and P_2 , but not under $\mathcal{P} := \text{conv}\{P_1, P_2\}$.⁹

A similar statement can be made for robust no-arbitrage, but under the extra condition that \mathcal{P} has finite cardinality. Suppose $\text{NA}^r(P)$ holds $\forall P \in \mathcal{P}$ with $|\mathcal{P}| < \infty$. Then for each P , there is a less frictional market $[\underline{S}^P, \overline{S}^P] \subseteq \text{ri}[\underline{S}, \overline{S}]$ satisfying $\text{NA}(P)$. Define $\overline{S}' := \sup_{P \in \mathcal{P}} \overline{S}^P$ and $\underline{S}' := \inf_{P \in \mathcal{P}} \underline{S}^P$. Since \mathcal{P} is finite, it is easy to check that the $[\underline{S}', \overline{S}']$ -market has the desired measurability, lies in the relative interior of the original market, and satisfies $\text{NA}(\mathcal{P})$.

⁸“ri” stands for relative interior.

⁹“conv” stands for convex hull.

Definition 2.3.3. A pair (Q, \tilde{S}) is called a consistent price system (CPS) (resp. strictly consistent price system (SCPS)) if \tilde{S} is a Q -martingale and $\tilde{S}_t \in [\underline{S}_t, \overline{S}_t]$ (resp. $\tilde{S}_t \in ri[\underline{S}_t, \overline{S}_t]$) Q -a.s.. Denote the set of all consistent price systems (resp. strictly consistent price systems) by \mathcal{Z} (resp. \mathcal{Z}^s)

When a market \mathbf{M} other than $[\underline{S}, \overline{S}]$ is in discussion, we shall write $\mathcal{Z}_{\mathbf{M}}, \mathcal{Z}_{\mathbf{M}}^s$ to indicate the underlying market. We now state a continuity assumption under which our main theorems hold. Recall that for each $\omega \in \Omega_t$, $\mathcal{P}_t(\omega) \subseteq \mathfrak{P}(\Omega_1)$ is the collection of models for the $(t+1)$ -th period starting at position ω . The support of $\mathcal{P}_t(\omega)$, denoted by $\text{supp}\mathcal{P}_t(\omega)$, is the smallest closed set in Ω_1 whose complement is $\mathcal{P}_t(\omega)$ -polar. This gives us a set-valued map $\omega \mapsto \text{supp}\mathcal{P}_t(\omega)$.

Assumption 2.3.4. *For each $t \in \{0, \dots, T-1\}$, $\underline{S}_{t+1}(\cdot), \overline{S}_{t+1}(\cdot)$ are continuous functions, and $\text{supp}\mathcal{P}_t(\cdot)$ is continuous as a set-valued map.¹⁰*

The necessity of this assumption is briefly discussed in Section 2.3.3. It holds trivially when Ω_1 is a discrete space. It also covers many examples in the non-dominated case.

Example 2.3.5. Let $\Omega_t = \mathbb{R}_{\geq 0}^t$ be equipped with the uniform norm. Let $S_0 > 0$ and $S_t(\omega) = \omega_t$ (the t -th component of ω) for $t \geq 1$, $\omega \in \Omega_t$, representing the mid-price of the stock. Then $\underline{S}_t := (1 - \kappa)S_t$, $\overline{S}_t = (1 + \kappa)S_t$ for some positive constant κ are continuous for each t . (1) $\mathcal{P}_t \equiv \mathfrak{P}(\Omega_1)$, i.e. model-free. $\text{supp}\mathcal{P}_t \equiv \Omega_1$ is obviously a continuous set-valued map. (2) \mathcal{P}_t consists of all probability measures on Ω_1 such that S_{t+1}/S_t lies inside a given interval $[a, b]$, i.e. there is uncertainty in the log-increment of the mid-price. In this case $\text{supp}\mathcal{P}_t(\omega) = [a\omega_t, b\omega_t]$, which is continuous in ω .

¹⁰See Appendix A.2 for the definition of continuity of a set-valued map.

The main theorems of this section are given below.

Theorem 2.3.6. *Under Assumption 2.3.4, the following are equivalent:*

(i) $NA(\mathcal{P})$ holds.

(ii) $\forall P \in \mathcal{P}, \exists (Q, \tilde{S}) \in \mathcal{Z}$ such that $P \ll Q \lll \mathcal{P}$.¹¹

Theorem 2.3.7. *Under Assumption 2.3.4, the following are equivalent:*

(i) $NA^r(\mathcal{P})$ hold.

(ii) There exist bid-ask price processes $[\underline{S}', \overline{S}'] \subseteq ri[\underline{S}, \overline{S}]$, continuous for each t , and satisfying

$$(2.14) \quad \forall P \in \mathcal{P}, \exists (Q, \tilde{S}) \in \mathcal{Z}' \subseteq \mathcal{Z}^s \text{ such that } P \ll Q \lll \mathcal{P},$$

where $\mathcal{Z}' = \mathcal{Z}_{[\underline{S}', \overline{S}']}$.

Once we prove Theorem 2.3.6, the proof of Theorem 2.3.7 is easy. So we focus on the former and skip the latter in the one-period case. [25, Theorem 3.1] and [71, Theorem 1.7] suggest that it may be natural to formulate (ii) of Theorem 2.3.7 as:

$$(2.15) \quad \forall P \in \mathcal{P}, \exists (Q, \tilde{S}) \in \mathcal{Z}^s \text{ such that } P \ll Q \lll \mathcal{P},$$

which is weaker than condition (2.14) and implies $NA(\mathcal{P})$. Under the weaker condition, we get a pair (Q^P, \tilde{S}^P) for every P . When \mathcal{P} is finite, in particular, a singleton, the two conditions are equivalent, since one can construct $\overline{S}', \underline{S}'$ as the pointwise maximum and minimum of all the \tilde{S}^P 's, respectively. However, when \mathcal{P} has infinitely many elements, it is not clear how to construct a less frictional market that is arbitrage-free with respect to \mathcal{P} : taking pointwise supremum and infimum does

¹¹The notation $Q \lll \mathcal{P}$ is taken from [25]. It means $Q \ll P$ for some $P \in \mathcal{P}$.

not necessarily produce a market that lies in the relative interior of the original one. Finally, let \mathcal{Q} (resp. \mathcal{Q}') be the collection of the first components of \mathcal{Z} (resp. \mathcal{Z}') that are “strongly” absolutely continuous with respect to \mathcal{P} . We remark that Theorem 2.3.6(ii) is equivalent to saying \mathcal{P}, \mathcal{Q} are equivalent in terms of polar sets, and Theorem 2.3.7(ii) says $\mathcal{P}, \mathcal{Q}'$ are equivalent in terms of polar sets. When \mathcal{P} is a singleton, we recover the classical result of the existence of an equivalent measure.

2.3.2 The building block: the one-period case

In this subsection, we prove the FTAP for a one-period market. To prepare for multi-period case, we also discuss how to construct martingales with certain prescribed initial values. Throughout this section, we do not impose the continuity assumption, since this assumption is required in carrying out the backward recursion, which clearly is unnecessary when there is only one period.

Let (Ω, \mathcal{F}) be a measurable space with filtration $(\mathcal{F}_0, \mathcal{F}_1)$ and $\mathcal{F}_0 = \{0, \Omega\}$. Let $\mathcal{P} \subseteq \mathfrak{P}(\Omega)$ be a nonempty convex set. The bid and ask price processes of the stock are given by constants $\underline{S}_0, \bar{S}_0$ and \mathcal{F}_1 -measurable random variables $\underline{S}_1, \bar{S}_1$, respectively. For this one-period market, we write ΔX for the difference $X_1 - X_0$ of any process X . Finally, we note that $\text{NA}(\mathcal{P})$ for this one-period market can be stated in the following equivalent form: $\forall y \in \mathbb{R}, y^+(\underline{S}_1 - \bar{S}_0) - y^-(\bar{S}_1 - \underline{S}_0) \geq 0$ \mathcal{P} -q.s. implies $y^+(\underline{S}_1 - \bar{S}_0) - y^-(\bar{S}_1 - \underline{S}_0) = 0$ \mathcal{P} -q.s..

For each $P \in \mathcal{P}$, define

$$(2.16) \quad \Theta_P := \{R \in \mathfrak{P}(\Omega) : P \ll R \lll \mathcal{P}, E^R[|\bar{S}_1 - \underline{S}_0| + |\underline{S}_1 - \bar{S}_0|] < \infty\}.$$

Θ_P is nonempty by Lemma A.3.1.

Lemma 2.3.8. *Let $\text{NA}(\mathcal{P})$ hold. Define \tilde{S}^λ by $\tilde{S}_0^\lambda := \lambda \underline{S}_0 + (1 - \lambda) \bar{S}_0$, $\tilde{S}_1^\lambda :=$*

$\lambda\bar{S}_1 + (1 - \lambda)\underline{S}_1$. Then $\forall P \in \mathcal{P}$ and Θ_P defined by (2.16), we have

$$(2.17) \quad 0 \in \{E^R[\Delta\tilde{S}^\lambda] : R \in \Theta_P, \lambda \in [0, 1]\}.$$

If in addition, $\exists P_1, P_2 \in \mathcal{P}$ such that $P_1(\underline{S}_1 - \bar{S}_0 < 0) > 0$ and $P_2(\bar{S}_1 - \underline{S}_0 > 0) > 0$, then (2.17) holds with $\lambda \in (0, 1)$.

Proof. Let $P \in \mathcal{P}$. We consider three cases:

Case 1. $\exists R \in \Theta_P$ with $E^R[\underline{S}_1 - \bar{S}_0] > 0$. We claim that $\text{NA}(\mathcal{P})$ implies $\exists R' \in \Theta_P$ with $E^{R'}[\underline{S}_1 - \bar{S}_0] < 0$. To see this, first observe that we cannot have $\underline{S}_1 - \bar{S}_0 \geq 0$ \mathcal{P} -q.s. because $\text{NA}(\mathcal{P})$ would then imply $\underline{S}_1 - \bar{S}_0 = 0$ \mathcal{P} -q.s. and therefore R -a.s., contradicting $E^R[\underline{S}_1 - \bar{S}_0] > 0$. So the set $A := \{\underline{S}_1 - \bar{S}_0 < 0\}$ satisfies $R_1(A) > 0$ for some $R_1 \in \mathcal{P}$. Similar to the first paragraph on page 13 of [25], we define $R_2 := (R_1 + P)/2$, use Lemma A.3.1 to replace R_2 by $R_3 \sim R_2$ such that $R_3 \in \Theta_P$, and further replace R_3 by $R' \sim R_3$ defined by $dR'/dR_3 := (1_A + \epsilon)/E^{R_3}[1_A + \epsilon]$. It can be checked that $E^{R'}[\underline{S}_1 - \bar{S}_0] < 0$ for ϵ small enough. So with $\lambda = 0$, we have found measures $R, R' \in \Theta_P$ such that $E^R[\Delta\tilde{S}^\lambda] > 0$ and $E^{R'}[\Delta\tilde{S}^\lambda] < 0$. In fact, any λ sufficiently close to zero will work since $\lambda \mapsto E^{R_0}[\Delta\tilde{S}^\lambda]$ is continuous for any $R_0 \in \Theta_P$. By taking suitable convex combination of R and R' , we can find a measure $Q \in \Theta_P$ satisfying $E^Q[\Delta\tilde{S}^\lambda] = 0$.

Case 2. $\exists R \in \Theta_P$ with $E^R[\bar{S}_1 - \underline{S}_0] < 0$. Similar to case 1, $\text{NA}(\mathcal{P})$ implies $\exists R' \in \Theta_P$ with $E^{R'}[\bar{S}_1 - \underline{S}_0] > 0$. We can pick any $\lambda \in [0, 1]$ sufficiently close to 1, and a suitable convex combination of R and R' , denoted by Q , such that $E^Q[\Delta\tilde{S}^\lambda] = 0$.

Case 3. $\forall R \in \Theta_P$, $E^R[\underline{S}_1 - \bar{S}_0] \leq 0$ and $E^R[\bar{S}_1 - \underline{S}_0] \geq 0$. Pick any $R \in \Theta_P$. If one of the inequalities is an equality, then we are done: $\lambda = 0$ or 1 will do. If both

inequalities are strict, then we can always find a $\lambda \in (0, 1)$ such that

$$E^R[\Delta \tilde{S}^\lambda] = \lambda E^R[\bar{S}_1 - \underline{S}_0] + (1 - \lambda) E^R[\underline{S}_1 - \bar{S}_0] = 0.$$

If $\exists P_1, P_2 \in \mathcal{P}$ such that $P_1(\underline{S}_1 - \bar{S}_0 < 0) > 0, P_2(\bar{S}_1 - \underline{S}_0 > 0) > 0$, then for some $Q \in \Theta_P$, both inequalities can be made strict: we can find $R_1, R_2 \in \Theta_P$ such that $E^{R_1}[\underline{S}_1 - \bar{S}_0] < 0$ and $E^{R_2}[\bar{S}_1 - \underline{S}_0] > 0$ by a construction similar to that in case 1. We then take $Q := (R_1 + R_2)/2$. Such a Q satisfies $E^Q[\underline{S}_1 - \bar{S}_0] < 0$ and $E^Q[\bar{S}_1 - \underline{S}_0] > 0$. \square

Given $P \in \mathcal{P}$, (2.17) immediately gives the existence of a CPS (Q, \tilde{S}) with $P \lll Q \lll \mathcal{P}$, which is the nontrivial implication of the FTAP. The additional claim in Lemma 2.3.8 says the only case when we fail to have an SCPS is $\underline{S}_1 \geq \bar{S}_0$ or $\bar{S}_1 \leq \underline{S}_0$ \mathcal{P} -q.s.. Under $\text{NA}(\mathcal{P})$, this is only possible if $\underline{S}_1 = \bar{S}_0$ or $\bar{S}_1 = \underline{S}_0$ \mathcal{P} -q.s.. Lemma 2.3.8 also has the following important implications for the multi-period case.

Remark 2.3.9. A consistent price process of the form $\tilde{S} = \tilde{S}^\lambda$ inherits the measurability of the given bid-ask prices. In multi-period case, the measurability also depends on the measurability of λ . By focusing on consistent price processes of this special form, for each single period, we can reduce the problem of finding a measurable function \tilde{S}_1 to finding a constant weight λ , which will be useful when pasting single-period prices together using measurable selection in the multi-period case. Otherwise we could encounter the problem of having $\tilde{S}_{t+1}(\omega, \omega')$ being \mathcal{F}_t measurable in ω and \mathcal{F}_1 -measurable in ω' , but not necessarily \mathcal{F}_{t+1} -measurable in (ω, ω') .

Proposition 2.3.10. *In a one-period market, the following are equivalent:*

- (i) $\text{NA}(\mathcal{P})$ holds.
- (ii) $\forall P \in \mathcal{P}, \exists (Q, \tilde{S}) \in \mathcal{Z}$ such that $P \lll Q \lll \mathcal{P}$.

Proof. (i) \Rightarrow (ii) follows from Lemma 2.3.8. (ii) \Rightarrow (i) is a simplified version of its multi-period counterparts (see the proof of Theorem 2.3.6). \square

Recall that when we go to the multi-period case, we cannot directly paste two single-period CPSs, but to first make sure the starting point of the current-period martingale matches the terminal point of its parent. In other words, we are interested in constructing martingales with certain prescribed initial values. Proposition 2.3.11 gives the set of starting points that admits a martingale extension.

For a random variable $S : \Omega \rightarrow \mathbb{R}$ and a nonempty family \mathcal{R} of probability measures on Ω , $\text{supp}_{\mathcal{R}}S$ denotes the smallest closed set $A \subseteq \mathbb{R}$ such that $P(S \in A) = 1 \forall P \in \mathcal{R}$.

Proposition 2.3.11. *Let $s \in \text{ri}[\inf \text{supp}_{\mathcal{P}}\underline{S}_1, \sup \text{supp}_{\mathcal{P}}\bar{S}_1]$. Then $NA(\mathcal{P})$ holds for the market $\{s, [\underline{S}_1, \bar{S}_1]\}$. And $\forall P \in \mathcal{P}$, $\exists(Q, \tilde{S}) \in \mathcal{Z}$ such that $P \ll Q \lll \mathcal{P}$, $\tilde{S}_0 = s$, and $\tilde{S}_1 = \lambda \bar{S}_1 + (1 - \lambda)\underline{S}_1$ for some $\lambda \in (0, 1)$.*

Proof. If $\text{supp}_{\mathcal{P}}\underline{S}_1 = \text{supp}_{\mathcal{P}}\bar{S}_1 = \{s\}$, then $NA(\mathcal{P})$ holds trivially and $Q := P, \tilde{S}_0 := s, \tilde{S}_1 := (\underline{S}_1 + \bar{S}_1)/2$ is the desired CPS. Suppose $s \in (\inf \text{supp}_{\mathcal{P}}\underline{S}_1, \sup \text{supp}_{\mathcal{P}}\bar{S}_1) \neq \emptyset$. We can find $x \in \text{supp}_{\mathcal{P}}\underline{S}_1$ and $y \in \text{supp}_{\mathcal{P}}\bar{S}_1$ such that $x < s < y$. By definition of support, $\exists P_1, P_2 \in \mathcal{P}$ satisfying $P_1(\underline{S}_1 < s) > 0$ and $P_2(\bar{S}_1 > s) > 0$. We now show the market $\{s, [\underline{S}_1, \bar{S}_1]\}$ satisfies $NA(\mathcal{P})$. Let $y \in \mathbb{R}$ satisfy $y^+(\underline{S}_1 - s) - y^-(\bar{S}_1 - s) \geq 0$ \mathcal{P} -q.s.. If $y > 0$, then we must have $\underline{S}_1 \geq s$ \mathcal{P} -q.s., contradicting the fact that $P_1(\underline{S}_1 < s) > 0$. If $y < 0$, then we must have $\bar{S}_1 \leq s$ \mathcal{P} -q.s., contradicting the fact that $P_2(\bar{S}_1 > s) > 0$. Therefore, the only possibility is $y = 0$, thus $y^+(\underline{S}_1 - s) - y^-(\bar{S}_1 - s) = 0$ \mathcal{P} -q.s.. Applying Lemma 2.3.8 to the market $\{s, [\underline{S}_1, \bar{S}_1]\}$ yields the desired CPS and a $\lambda \in (0, 1)$. \square

2.3.3 The multi-period case

In this subsection, we prove Theorem 2.3.6 for a multi-period market through a backward-forward scheme. Back to the setup in introduction, the set \mathcal{P} is defined as the product, in the sense of (2.1.1), of the nonempty convex sets $\mathcal{P}_t(\cdot)$ which have analytic graphs, and $\underline{S}_t, \bar{S}_t$ are positive, \mathcal{F}_t -measurable. For a map f on Ω_{t+1} , we will often see it as a map on $\Omega_t \times \Omega_1$ and write $f = f(\omega, \omega')$. Throughout this section, Assumption 2.3.4 is in force. That is, we assume $\underline{S}_t(\cdot), \bar{S}_t(\cdot)$ and $\text{supp}\mathcal{P}_t(\cdot)$ are continuous. The reason for the extra assumption is that we wish to have a property for price processes that can be preserved under the backward recursion (2.18). Both Borel and universal measurability are lost under this backward scheme (see Remark 2.3.13). This problem does not exist in market without transaction cost since there is no need to redefine the stock price and it is enough to assume the stock price is Borel, nor does it matter when there is a dominating measure P , since we can always modify a universally measurable map on a P -null set to make it Borel. Relaxation of the continuity restriction is left for future research.

Define processes X, Y recursively by $X_T = \underline{S}_T, Y_T = \bar{S}_T$ and

$$(2.18) \quad X_t(\omega) := (\inf \text{supp}_{\mathcal{P}_t(\omega)} X_{t+1}(\omega, \cdot)) \vee \underline{S}_t(\omega) \wedge \bar{S}_t(\omega),$$

$$Y_t(\omega) := (\sup \text{supp}_{\mathcal{P}_t(\omega)} Y_{t+1}(\omega, \cdot)) \wedge \bar{S}_t(\omega) \vee \underline{S}_t(\omega)$$

for $t = T - 1, \dots, 0$.

Lemma 2.3.12. *For each t , $X_t, Y_t : \Omega_t \rightarrow \mathbb{R}$ are continuous.*

Proof. X_T, Y_T are continuous by assumption. Suppose X_{t+1}, Y_{t+1} are continuous, we deduce the continuity of X_t, Y_t . Lemma A.1.3 yields the nice representation

$$(2.19) \quad \begin{aligned} \text{supp}_{\mathcal{P}_t(\omega)} X_{t+1}(\omega, \cdot) &= \overline{X_{t+1}(\omega, \text{supp}(\mathcal{P}_t(\omega)))}, \\ \text{supp}_{\mathcal{P}_t(\omega)} Y_{t+1}(\omega, \cdot) &= \overline{Y_{t+1}(\omega, \text{supp}(\mathcal{P}_t(\omega)))}. \end{aligned}$$

Since $\omega \mapsto \text{supp}\mathcal{P}_t(\omega)$ is continuous, it is easy to check the set-valued map $\omega \mapsto \{\omega\} \times \text{supp}\mathcal{P}_t(\omega)$ is also continuous. Composing with the continuous function $(\omega, \omega') \mapsto X_{t+1}(\omega, \omega') \vee \underline{S}_t(\omega) \wedge \overline{S}_t(\omega)$ and then taking closure, we get by [3, Lemma 17.22, Theorem 17.23] a continuous map with non-empty compact values:

$$\overline{\{X_{t+1}(\omega, \omega') \vee \underline{S}_t(\omega) \wedge \overline{S}_t(\omega) : \omega' \in \text{supp}(\mathcal{P}_t(\omega))\}} := \Phi(\omega)$$

The Berge Maximum Theorem [3, Theorem 17.31] then implies that the value function $\inf_{y \in \Phi(\omega)} y = \inf \Phi(\omega)$ is continuous. It remains to notice that

$$\inf \Phi(\omega) = \inf \overline{X_{t+1}(\omega, \text{supp}(\mathcal{P}_t(\omega))) \vee \underline{S}_t(\omega) \wedge \overline{S}_t(\omega)} = X_t.$$

A symmetric argument gives the continuity of Y_t . □

Remark 2.3.13. If X_{t+1}, Y_{t+1} are Borel measurable, we can show X_t, Y_t are universally measurable. Indeed, from [25, Lemma 4.3], we know the closed-valued maps $\text{supp}_{\mathcal{P}_t(\omega)} X_{t+1}(\omega, \cdot)$, $\text{supp}_{\mathcal{P}_t(\omega)} Y_{t+1}(\omega, \cdot)$ are universally measurable. Castaing representation [3, Corollary 18.14] implies the infimum and supremum functions are also measurable. However, universal measurability is not preserved in the next iteration. See Remark 4.4 of [25] for a counter example.

Apart from preserving continuity, the recursively defined $[X, Y]$ -market has two nice properties. First, its spread is not too wide: at least all points in the interior of $[X_t, Y_t]$ admits a martingale extension to the next period \mathcal{P} -q.s., although there are delicate issues when the point lies on the boundary of the spread. Second, its spread is not too narrow either, in the sense that it still satisfies $\text{NA}(\mathcal{P})$ when the original market does. In summary, other than requiring a strong continuity assumption in its construction, this new market fits our needs perfectly. The general idea of proving the nontrivial implication of multi-period FTAP is to replace the original market by

the modified market $[X, Y]$, and do martingale extension in the modified market. Interior extension is not too hard in view of Proposition 2.3.11; the challenging part is boundary extension. It turns out that boundary extension is possible if we avoid hitting boundaries as much as we can from the beginning.

Before proving the main theorem, we need three crucial lemmas.

Lemma 2.3.14. *Let $NA(\mathcal{P})$ hold for the original market $[\underline{S}, \bar{S}]$. Then $NA(\mathcal{P})$ also holds for the modified market $[X, Y]$.*

Proof. We prove by backward induction. Suppose $NA(\mathcal{P})$ holds for the market $\mathbf{M}^{t+1} := \{[\underline{S}_r, \bar{S}_r]_{r=0, \dots, t}, [X_r, Y_r]_{r=t+1, \dots, T}\}$. We show $NA(\mathcal{P})$ holds for the market $\mathbf{M}^t := \{[\underline{S}_r, \bar{S}_r]_{r=0, \dots, t-1}, [X_r, Y_r]_{r=t, \dots, T}\}$.

Let ϕ be a self-financing portfolio process in the market \mathbf{M}^t with $\phi_0 = 0$ and $\phi_{T+1} \geq 0$ \mathcal{P} -q.s.. Consider another portfolio process defined by $\eta_r := \phi_r \forall r = 0, \dots, t$,

$$\begin{aligned} \Delta \eta_{t+1}^1 &:= 1_{\{Y_t = \bar{S}_t\}} (\Delta \phi_{t+1}^1)^+ - 1_{\{X_t = \underline{S}_t\}} (\Delta \phi_{t+1}^1)^-, \\ \Delta \eta_{t+2}^1 &:= \Delta \phi_{t+2}^1 + 1_{\{Y_t \neq \bar{S}_t\}} (\Delta \phi_{t+1}^1)^+ - 1_{\{X_t \neq \underline{S}_t\}} (\Delta \phi_{t+1}^1)^-, \\ \Delta \eta_{r+1}^1 &:= \Delta \phi_{r+1}^1, \quad r = t+2, \dots, T, \\ \Delta \eta_{r+1}^0 &:= -(\Delta \eta_{r+1}^1)^+ X_r + (\Delta \eta_{r+1}^1)^- Y_r, \quad r = t, \dots, T. \end{aligned}$$

That is, we follow ϕ up to time $t-1$, stick to its stock position whenever the transaction at time t can be carried out in the market \mathbf{M}^{t+1} , and postpone the transaction to time $t+1$ if it is not admissible in the market \mathbf{M}^{t+1} , and follow the stock position of ϕ again afterwards. Clearly, η is predictable, self-financing in the market \mathbf{M}^{t+1} , and $\eta_{T+1}^1 = \phi_{T+1}^1$. We want to show $\eta_{T+1}^0 \geq \phi_{T+1}^0$ \mathcal{P} -q.s.. It suffices to show $\Delta \eta_{t+1}^0 + \Delta \eta_{t+2}^0 \geq \Delta \phi_{t+1}^0 + \Delta \phi_{t+2}^0$. During the $(t+1)$ -th and $(t+2)$ -th periods, η and ϕ are trading the same total number of shares, just at different times. So we

only need to check that η faces a trading price as favorable as, if not more favorable than the one faced by ϕ . By our construction of X_t , when $X_t(\omega) \neq \underline{S}_t(\omega)$, we must have $X_t(\omega) \leq X_{t+1}(\omega, \cdot)$ $\mathcal{P}_t(\omega)$ -q.s.. Fubini theorem implies the \mathcal{F}_{t+1} -measurable set $\{X_t \neq \underline{S}_t\} \cap \{X_t > X_{t+1}\}$ is \mathcal{P} -polar. Similarly, $\{Y_t \neq \bar{S}_t\} \cap \{Y_t < Y_{t+1}\}$ is \mathcal{P} -polar. Therefore, η has price disadvantage only on a \mathcal{P} -polar set. \square

Lemma 2.3.15. *Let $[\underline{S}, \bar{S}]$ be any Borel market and let $t \in \{0, \dots, T-1\}$. Then*

$$N_t = \{\omega \in \Omega_t : NA(\mathcal{P}_t(\omega)) \text{ fails}\}$$

is universally measurable. If $NA(\mathcal{P})$ hold, then N_t is \mathcal{P} -polar.

Proof. Set $F^\omega(\cdot) := \underline{S}_{t+1}(\omega, \cdot) - \bar{S}_t(\omega)$ and $G^\omega(\cdot) := \bar{S}_{t+1}(\omega, \cdot) - \underline{S}_t(\omega)$. By Lemma 4.3 of [25], $\Lambda_F(\omega) := \text{supp}_{\mathcal{P}_t(\omega)}(F^\omega)$ and $\Lambda_G(\omega) := \text{supp}_{\mathcal{P}_t(\omega)}(G^\omega)$ are universally measurable. We claim that

$$(2.20) \quad N_t^c = \{\Lambda_F = \{0\}\} \cup \{\Lambda_G = \{0\}\} \cup \{\Lambda_F \cap \mathbb{R}_{<0} \neq \emptyset \text{ and } \Lambda_G \cap \mathbb{R}_{>0} \neq \emptyset\},$$

$$(2.21) \quad N_t = \{\Lambda_F \subseteq \mathbb{R}_{\geq 0} \text{ and } \Lambda_F \cap \mathbb{R}_{>0} \neq \emptyset\} \cup \{\Lambda_G \subseteq \mathbb{R}_{\leq 0} \text{ and } \Lambda_G \cap \mathbb{R}_{<0} \neq \emptyset\}.$$

Indeed, if $\Lambda_F(\omega) = \{0\}$ or $\Lambda_G(\omega) = \{0\}$, $\omega \in N_t^c$ trivially. If $\Lambda_F(\omega) \cap \mathbb{R}_{<0} \neq \emptyset$ and $\Lambda_G(\omega) \cap \mathbb{R}_{>0} \neq \emptyset$, let $y \in \mathbb{R}$ satisfy $y^+ F^\omega - y^- G^\omega \geq 0$ $\mathcal{P}_t(\omega)$ -q.s.. Suppose $y > 0$, then we get $F^\omega \geq 0$ $\mathcal{P}_t(\omega)$ -q.s., contradicting $\Lambda_F(\omega) \cap \mathbb{R}_{<0} \neq \emptyset$. Suppose $y < 0$, we get $G^\omega \leq 0$ $\mathcal{P}_t(\omega)$ -q.s., contradicting $\Lambda_G(\omega) \cap \mathbb{R}_{>0} \neq \emptyset$. So we must have $y = 0$, and consequently $\omega \in N_t^c$. Conversely, if $\Lambda_F(\omega) \subseteq \mathbb{R}_{\geq 0}$ and $\Lambda_F(\omega) \cap \mathbb{R}_{>0} \neq \emptyset$, then $F^\omega \geq 0$ $\mathcal{P}_t(\omega)$ -q.s. and $P(F^\omega > 0) > 0$ for some $P \in \mathcal{P}_t(\omega)$. In this case, any $y > 0$ is an arbitrage. Similarly, if $\Lambda_G(\omega) \subseteq \mathbb{R}_{\leq 0}$ and $\Lambda_G(\omega) \cap \mathbb{R}_{<0} \neq \emptyset$, then any $y < 0$ is an arbitrage. Universal measurability of N_t then follows from the universal measurability of Λ_F, Λ_G . We now show N_t is \mathcal{P} -polar.

We only show that the set $A := \{\Lambda_F \subseteq \mathbb{R}_{\geq 0} \text{ and } \Lambda_F \cap \mathbb{R}_{> 0} \neq \emptyset\}$ is \mathcal{P} -polar. The other set is treated similarly. Suppose $\exists P_* \in \mathcal{P}$ such that $P_*(A) > 0$. To produce a simple buy-and-sell arbitrage, we want a universally measurable selector from $\Xi(\omega) := \{P \in \mathcal{P}_t(\omega) : E^P[F^\omega] > 0\}$ for $\omega \in A$, which can be done if the set-valued map has an analytic graph. Define $\psi(\omega, P) := E^P[F^\omega]$. ψ is Borel measurable by [22, Propositions 7.25, 7.26, 7.29]. So $\text{graph}(\Xi) = \text{graph}(\mathcal{P}_t) \cap \{\psi > 0\}$ is analytic, and consequently, Ξ admits a universally measurable selector $P(\cdot)$ on $\{\Xi \neq \emptyset\} \supseteq A$ by the Jankov-von Neumann theorem (Theorem A.2.3). On the universally measurable set A^c , redefine P to be any universally measurable selector of \mathcal{P}_t . Let $\phi_{t+1}^1 = 1_A, \phi_r^1 = 0$ for all $r \neq t+1$, and $\phi_0^0 = 0, \Delta\phi_{r+1}^0 = -(\Delta\phi_{r+1}^1)^+ \bar{S}_r + (\Delta\phi_{r+1}^1)^- \underline{S}_r$ for $r = 0, \dots, T$. Then $\phi \in \mathcal{H}$, $\phi_{t+2}(\omega, \cdot) \geq 0$ $\mathcal{P}_t(\omega)$ -q.s. for all $\omega \in \Omega$, and $P(\omega)(\phi_{t+2}^0(\omega, \cdot) > 0) > 0$ for $\omega \in A$.

Since each measure in \mathcal{P} admits a decomposition of the form (2.1.1), Fubini's theorem easily implies $\phi_{T+1} = \phi_{t+2} \geq 0$ \mathcal{P} -q.s.. On the other hand, $P^* := P_*|_{\Omega_t} \otimes P \otimes \tilde{P}_{t+1} \otimes \dots \otimes \tilde{P}_{T-1}$ where \tilde{P}_r is any universally measurable selector of $\mathcal{P}_r, r = t+1, \dots, T-1$ is an element of \mathcal{P} satisfying $P^*(\phi_{T+1}^0 > 0) = P^*(\phi_{t+2}^0 > 0) > 0$. This violates $\text{NA}(\mathcal{P})$. So A must be \mathcal{P} -polar. \square

Lemma 2.3.16. *Let $[\underline{S}, \bar{S}]$ be any Borel market. Let $t \in \{0, \dots, T-1\}$ and $P(\cdot) : \Omega_t \rightarrow \mathfrak{P}(\Omega_1)$ be Borel. Given a Borel measurable function $\tilde{S}_t(\cdot) \in [\underline{S}_t(\cdot), \bar{S}_t(\cdot)]$, let*

$$(2.22)$$

$$\Xi_t(\omega) := \{(Q, \lambda, \hat{P}) \in \mathfrak{P}(\Omega_1) \times (0, 1) \times \mathcal{P}_t(\omega) : P(\omega) \ll Q \ll \hat{P}, E^Q[D^\lambda(\omega)] = 0\},$$

where $D^\lambda(\omega) := \lambda \bar{S}_{t+1}(\omega, \cdot) + (1-\lambda) \underline{S}_{t+1}(\omega, \cdot) - \tilde{S}_t(\omega)$. Then Ξ_t has an analytic graph and there exist \mathcal{F}_t -measurable maps $Q(\cdot), \lambda(\cdot), \hat{P}(\cdot)$ such that $(Q(\omega), \lambda(\omega), \hat{P}(\omega)) \in \Xi_t(\omega)$ if $\Xi_t(\omega) \neq \emptyset$.

Proof. The proof is almost the same as that of [25, Lemma 4.8]. So we shall be brief. We first show Ξ_t has an analytic graph. Let $\Psi(\omega) := \{(Q, \lambda) \in \mathfrak{P}(\Omega_1) \times (0, 1) : E^Q[D^\lambda(\omega)] = 0\}$. Since the function $(\omega, Q, \lambda) \mapsto E^Q[D^\lambda(\omega)]$ is Borel, Ψ has Borel graph. Let $\Phi(\omega) := \{(R, \hat{R}) \in \mathfrak{P}(\Omega_1) \times \mathfrak{P}(\Omega_1) : P(\omega) \ll R \ll \hat{R}\}$. Define $\phi(\omega, R, \hat{R}) := E^R[dP(\omega)/dR] + E^{\hat{R}}[dR/d\hat{R}]$ where we choose a version of the Radon-Nikodym derivatives (using absolutely continuous part) that are jointly Borel measurable as described in [25, Lemma 4.7] (see also [34, Theorem V.58] and the remark after it). [22, Propositions 7.26, 7.29] then imply ϕ is Borel. So $\text{graph}(\Phi) = \{\phi = 2\}$ is Borel. Hence, with minor abuse of notation, $\Xi_t(\omega) = (\Psi(\omega) \times \mathcal{P}_t(\omega)) \cap (\Phi(\omega) \times \mathbb{R})$ has an analytic graph. We can find a universally measurable selector $Q(\cdot), \lambda(\cdot), \hat{P}(\cdot)$ for Ξ on the universally measurable set $\{\Xi_t \neq \emptyset\}$. Outside this set, we simply define $Q(\cdot) = \hat{P}(\cdot) = P(\cdot)$ and λ to be any constant. \square

We are now ready to prove our main results.

Proof of Theorem 2.3.6. (i) \Rightarrow (ii): We first replace the original $[\underline{S}, \overline{S}]$ -market by the modified $[X, Y]$ -market which lies inside $[\underline{S}, \overline{S}]$, is still continuous (hence Borel) by Lemma 2.3.12 and satisfies NA(\mathcal{P}) by Lemma 2.3.14. It suffices to prove (ii) for the modified market because any CPS for the modified market is a CPS for the original market. Let us introduce and prove an auxiliary claim:

(ii') $\forall P \in \mathcal{P}, \exists(Q, \tilde{S}) \in \mathcal{Z}$ such that $P \ll Q \ll \mathcal{P}$ and $\tilde{S}_t = \tilde{S}_t^\lambda = (1 - \lambda_{t-1})X_t + \lambda_{t-1}Y_t, t = 1, \dots, T$ for some adapted process λ , valued in $[0, 1]$. Moreover, define a sequence of stopping times:

$$\begin{aligned} \tau_1^0 &:= \inf\{t \in [0, T-1] : \lambda_t = 0\}, \\ \sigma_n^0 &:= \inf\{t \in (\tau_n^0, T-1] : \lambda_t > 0\}, \\ \tau_{n+1}^0 &:= \inf\{t \in (\sigma_n^0, T-1] : \lambda_t = 0\}, \end{aligned}$$

with the convention that $\inf \emptyset = \infty$. Then \mathcal{P} -q.s. on the set $\{\tau_n^0 < \infty\}$, we have $X_{\tau_n^0} = Y_{\tau_n^0} = X_t$ for all $t \in [\tau_n^0, \sigma_n^0 \wedge T]$. Similarly, define

$$\begin{aligned}\sigma_1^1 &:= \inf\{t \in [0, T-1] : \lambda_t = 1\}, \\ \tau_n^1 &:= \inf\{t \in (\sigma_n^1, T-1] : \lambda_t < 1\}, \\ \sigma_{n+1}^1 &:= \inf\{t \in (\tau_n^1, T-1] : \lambda_t = 1\}.\end{aligned}$$

Then \mathcal{P} -q.s. on the set $\{\sigma_n^1 < \infty\}$, we have $X_{\sigma_n^1} = Y_{\sigma_n^1} = Y_t$ for all $t \in [\sigma_n^1, \tau_n^1 \wedge T]$.

We do induction on the number of periods in the market. When there is only one period, for any $P \in \mathcal{P}$, the existence of $(Q, \tilde{S}^\lambda) \in \mathcal{Z}$ with $P \ll Q \lll \mathcal{P}$ is due to Lemma 2.3.8. Moreover, we take $\lambda = 0$ when $\{X_0\} = \{Y_0\} = \text{supp}_{\mathcal{P}} X_1 \neq \text{supp}_{\mathcal{P}} Y_1$, and $\lambda = 1$ when $\{X_0\} = \{Y_0\} = \text{supp}_{\mathcal{P}} Y_1 \neq \text{supp}_{\mathcal{P}} X_1$. In all other cases (under $\text{NA}(\mathcal{P})$), Proposition 2.3.11 guarantees the existence of a $\lambda \in (0, 1)$. We can check that all the statements in (ii') are satisfied.

Now, suppose (i) implies (ii') for any market with $t-1$ periods ($t = 2, \dots, T$) and satisfies backward recursion (2.18). We will deduce the same property for such recursively defined market with t periods.

Let $\text{NA}(\mathcal{P})$ hold for the t -period market denoted by \mathbf{M} . Its submarket up to time $t-1$, denoted by \mathbf{M}' , satisfies $\text{NA}(\mathcal{P}')$ where

$$\mathcal{P}' = \{P_0 \otimes \dots \otimes P_{t-2} : \text{each } P_r \text{ is a universally measurable selector of } \mathcal{P}_r\}$$

is the set of possible models for the first $t-1$ periods. Let $P \in \mathcal{P}$ have decomposition $P = P|_{\Omega_{t-1}} \otimes P_{t-1}$. We can apply the induction hypothesis to obtain Q', λ, \tilde{S} described in (ii') up to time $t-1$ with $P|_{\Omega_{t-1}} \lll Q' \lll \hat{P}' \in \mathcal{P}'$. Our goal is to extend $Q', \lambda, \tilde{S}, \hat{P}'$ to the t -th period.

Step 1: We will show that the set

$$(2.23) \quad N_{\mathbf{M}_1} := \{\omega \in \Omega_{t-1} : \text{NA}(\mathcal{P}_{t-1}(\omega)) \text{ fails for the one-period market} \\ \{\tilde{S}_{t-1}(\omega), [X_t(\omega, \cdot), Y_t(\omega, \cdot)]\}\}$$

is universally measurable and \mathcal{P} -polar. Let

$$\Lambda_F(\omega) := \text{supp}_{\mathcal{P}_{t-1}(\omega)} X_t(\omega, \cdot) - \tilde{S}_{t-1}(\omega),$$

$$\Lambda_G(\omega) := \text{supp}_{\mathcal{P}_{t-1}(\omega)} Y_t(\omega, \cdot) - \tilde{S}_{t-1}(\omega).$$

These sets are universally measurable by [25, Lemma 4.3]. Equation (2.21) gives

$$(2.24) \quad N_{\mathbf{M}_1} = \{\Lambda_F \subseteq \mathbb{R}_{\geq 0} \text{ and } \Lambda_F \cap \mathbb{R}_{>0} \neq \emptyset\} \cup \{\Lambda_G \subseteq \mathbb{R}_{\leq 0} \text{ and } \Lambda_G \cap \mathbb{R}_{<0} \neq \emptyset\},$$

which implies $N_{\mathbf{M}_1}$ is universally measurable. It remains to show $N_{\mathbf{M}_1}$ is \mathcal{P} -polar.

Since the market $[X, Y]$ satisfies $NA(\mathcal{P})$, Lemma 2.3.15 implies that $N_{t-1} := \{\omega : \text{NA}(\mathcal{P}_{t-1}(\omega)) \text{ fails for the one-period market } \{[X_{t-1}(\omega), Y_{t-1}(\omega)], [X_t(\omega, \cdot), Y_t(\omega, \cdot)]\}\}$ is universally measurable and \mathcal{P} -polar. So it suffices to show the set $N_{\mathbf{M}_1} \cap N_{R-1}^c$ is \mathcal{P} -polar. For $\omega \in N_{t-1}^c$, we see from (2.20) and the definition of X_{t-1}, Y_{t-1} that

$$[X_{t-1}(\omega), Y_{t-1}(\omega)] \subseteq [\inf \text{supp}_{\mathcal{P}_{t-1}(\omega)} X_t(\omega, \cdot), \sup \text{supp}_{\mathcal{P}_{t-1}(\omega)} Y_t(\omega, \cdot)].$$

Observe that $N_{t-1}^c \cap \{\Lambda_F \subseteq \mathbb{R}_{\geq 0} \text{ and } \Lambda_F \cap \mathbb{R}_{>0} \neq \emptyset\} \subseteq \{\lambda_{t-2} = 0\}$ because if $\lambda_{t-2}(\omega|_{\Omega_{t-2}}) > 0$ (here $\omega \in \Omega_{t-1}$ and $\omega|_{\Omega_{t-2}}$ denotes the first $t-2$ components of ω), then either $\tilde{S}_{t-1}(\omega) > X_{t-1}(\omega) \geq \inf \text{supp}_{\mathcal{P}_{t-1}(\omega)} X_t(\omega, \cdot)$ which would imply $\Lambda_F(\omega) \cap \mathbb{R}_{<0} \neq \emptyset$, or $X_{t-1}(\omega) = Y_{t-1}(\omega)$ in which case if $\Lambda_F(\omega) \subseteq \mathbb{R}_{\geq 0}$ and $\Lambda_F(\omega) \cap \mathbb{R}_{>0} \neq \emptyset$, then $\omega \in N_{t-1}$. Similarly, $N_{t-1}^c \cap \{\Lambda_G \subseteq \mathbb{R}_{\leq 0} \text{ and } \Lambda_G \cap \mathbb{R}_{<0} \neq \emptyset\} \subseteq \{\lambda_{t-2} = 1\}$. So to show $N_{\mathbf{M}_1} \cap N_{t-1}^c$ is \mathcal{P} -polar, it suffices to show the \mathcal{F}_{t-1} -measurable sets

$$A := \{\Lambda_F \subseteq \mathbb{R}_{\geq 0} \text{ and } \Lambda_F \cap \mathbb{R}_{>0} \neq \emptyset\} \cap \{\lambda_{t-2} = 0\},$$

$$B := \{\Lambda_G \subseteq \mathbb{R}_{\leq 0} \text{ and } \Lambda_G \cap \mathbb{R}_{<0} \neq \emptyset\} \cap \{\lambda_{t-2} = 1\}.$$

are \mathcal{P} -polar. We shall focus on set A ; the other case is similar.

Suppose on the contrary, $P_*(A) > 0$ for some $P_* \in \mathcal{P}$. Define stopping times $\tilde{\tau}_n, \tilde{\sigma}_n$ by

$$\begin{aligned}\tilde{\tau}_1^0 &:= \inf\{r \in [0, t-2] : \lambda_r = 0\}, \\ \tilde{\sigma}_n^0 &:= \inf\{r \in (\tilde{\tau}_n^0, t-2] : \lambda_r > 0\}, \\ \tilde{\tau}_{n+1}^0 &:= \inf\{r \in (\tilde{\sigma}_n^0, r-2] : \lambda_r = 0\}.\end{aligned}$$

Induction hypothesis implies that \mathcal{P}' -q.s. on $\{\tilde{\tau}_n^0 < \infty\}$, we have

$$(2.25) \quad X_{\tilde{\tau}_n^0} = Y_{\tilde{\tau}_n^0} = X_r \quad \forall r \in [\tilde{\tau}_n^0, \tilde{\sigma}_n^0 \wedge (t-1)]$$

We now construct an arbitrage strategy. Define $\phi_0 := 0$. For $r = 1, \dots, t$,

$$\phi_{r+1}^1 := \sum_n 1_{\{\tilde{\tau}_n^0 \leq r < \tilde{\sigma}_n^0\}} - 1_{\{\tilde{\tau}_n^0 < \tilde{\sigma}_n^0 = \infty\} \cap A^c \cap \{r=t-1\}} - 1_{\{\tilde{\tau}_n^0 < \tilde{\sigma}_n^0 = \infty\} \cap A \cap \{r=t\}},$$

and

$$\Delta \phi_{r+1}^0 = -(\Delta \phi_{r+1}^1)^+ X_r - (\Delta \phi_{r+1}^1)^- Y_r.$$

Then ϕ is predictable and self-financing in market \mathbf{M} . Moreover, $\phi_{t+1} \geq 0$ \mathcal{P} -q.s.. Indeed, on the set $\{\tilde{\tau}_1^0 = \infty\}$, no trade occurs. On the set $\{\tilde{\tau}_m^0 < \tilde{\sigma}_m^0 = \infty\}$ for some $m \geq 1$, the strategy is to repeatedly buy one share at time $\tilde{\tau}_n^0$ and sell it at time $\tilde{\sigma}_n^0$ for all $n < m$. After that, buy one share at $\tilde{\tau}_m^0$ and close our stock position at time $t-1$ if A is not observed, and at time t if A is observed. In the case where A is not observed, (2.25) implies the selling price of every holding period is \mathcal{P}' -q.s. (hence also \mathcal{P} -q.s. since all trades occur on or before time $t-1$) the same as the buying price of that holding period. So we end up in zero position. In the case where A is observed, (2.25) again implies \mathcal{P}' -q.s. perfect cancellation before the last holding period; in the last holding period, we buy a share at time $\tilde{\tau}_m^0$, and sell it at time t

at the price X_t which is \mathcal{P} -q.s. larger than or equal to $\tilde{S}_{t-1} = X_{t-1}$ by the definition of A . (2.25) then implies $X_t \geq Y_{\tilde{\tau}_m^0}$ \mathcal{P} -q.s. on $\{\tilde{\tau}_m^0 < \tilde{\sigma}_m^0 = \infty\} \cap A$. So we can close our position without loss. On the set $\{\tilde{\sigma}_m^0 < \tilde{\tau}_{m+1}^0 = \infty\}$ for some $m \geq 1$, all trades happen on or before time $t - 2$ and we have \mathcal{P}' -q.s. (hence \mathcal{P} -q.s.) perfect cancellation.

To create an arbitrage opportunity, it remains to construct a measure P^* under which $\{\phi_{t+1}^0 > 0\} = A \cap \{X_t > X_{t-1}\}$ has positive measure (notice that $\lambda_{t-2} = 0$ implies $\exists m \geq 1$ such that $\tilde{\tau}_m^0 < \tilde{\sigma}_m^0 = \infty$).

Let

$$\Xi(\omega) := \{R \in \mathcal{P}_{t-1}(\omega) : E^R[X_t(\omega, \cdot) - X_{t-1}(\omega)] > 0\}.$$

By [22, Propositions 7.25, 7.26, 7.29], the map defined by $\psi(\omega, R) := E^R[X_t(\omega, \cdot) - X_{t-1}(\omega)]$ is Borel measurable. Hence $\text{graph}(\Xi) = \text{graph}(\mathcal{P}_{t-1}) \cap \{\psi > 0\}$ is analytic. The Jankov-von Neumann theorem (Theorem A.2.3) implies Ξ admits a universally measurable selector $R(\cdot) \in \Xi(\cdot)$ on $\{\Xi \neq \emptyset\}$. Observe that $A \subseteq \{\Xi \neq \emptyset\}$. Outside A , we redefine R to be any universally measurable selector of \mathcal{P}_{t-1} . We have $R(\omega)(X_t(\omega, \cdot) - X_{t-1}(\omega) > 0) > 0$ on A . Define $P^* := P_*|_{\Omega_{t-1}} \otimes R$. Then $P^*(A) > 0$ and $P^*(\phi_{t+1}^0 > 0) > 0$. So ϕ is an arbitrage, contradicting $\text{NA}(\mathcal{P})$ for market \mathbf{M} . Therefore, A must be \mathcal{P} -polar. A similar argument shows that B is \mathcal{P} -polar. We conclude that $N_{\mathbf{M}_1} = (N_{\mathbf{M}_1} \cap N_{t-1}) \cup (N_{\mathbf{M}_1} \cap N_{t-1}^c)$ is \mathcal{P} -polar.

Step 2: We will now extend $Q', \lambda, \tilde{S}, \hat{P}'$ to the t -th period. In view of step 1, we only need to consider martingale extension on the \mathcal{P} -q.s. set $N_{\mathbf{M}_1}^c = \{\Lambda_F = \{0\}\} \cup \{\Lambda_G = \{0\}\} \cup \{\Lambda_F \cap \mathbb{R}_{<0} \neq \emptyset \text{ and } \Lambda_G \cap \mathbb{R}_{>0} \neq \emptyset\}$. On $N_{\mathbf{M}_1}$, we simply set $Q_{t-1} = \hat{P}_{t-1} = P_{t-1}$ and $\lambda_{t-1} = 1/2$. We perform the extension on the following universally measurable sets separately:

On $\{\Lambda_F = \{0\} \neq \Lambda_G\}$, set $Q_{t-1} = \hat{P}_{t-1} = P_{t-1}$, $\lambda_{t-1} = 0$.

On $\{\Lambda_G = \{0\} \neq \Lambda_F\}$, set $Q_{t-1} = \hat{P}_{t-1} = P_{t-1}$, $\lambda_{t-1} = 1$.

On $\{\Lambda_F = \Lambda_G = \{0\}\}$, set $Q_{t-1} = \hat{P}_{t-1} = P_{t-1}$, $\lambda_{t-1} = 1/2$.

On $\{\Lambda_F \cap \mathbb{R}_{<0} \neq \emptyset \text{ and } \Lambda_G \cap \mathbb{R}_{>0} \neq \emptyset\} := C$, we have

$$\inf \text{supp}_{\mathcal{P}_{t-1}(\omega)} X_t(\omega, \cdot) < \tilde{S}_{t-1}(\omega) < \sup \text{supp}_{\mathcal{P}_{t-1}(\omega)} Y_t(\omega, \cdot).$$

So the set $\Xi_{t-1}(\omega)$ defined by (2.22) with X, Y in place of \underline{S}, \bar{S} is nonempty by Proposition 2.3.11. To obtain a universally measurable selector, we first modify $P_{t-1}(\cdot)$ and $\tilde{S}_{t-1}(\cdot)$ on a \hat{P}' -nullset \hat{N} (hence Q' -null and $P|_{\Omega_{t-1}}$ -null) to make them Borel measurable [22, Lemma 7.27]. Denote the resulting Borel kernel and random variable by P_{t-1}^B and \tilde{S}_{t-1}^B . We can then use Lemma 2.3.16 to obtain universally measurable maps $Q_{t-1}^0(\cdot)$, $\lambda_{t-1}^0(\cdot)$, $\hat{P}_{t-1}^0(\cdot)$ such that $P_{t-1}^B(\omega) \ll Q_{t-1}^0(\omega) \ll \hat{P}_{t-1}^0(\omega)$, and if $\omega \in C \setminus \hat{N}$, then $\lambda_{t-1}^0(\omega) \in (0, 1)$, $\hat{P}_{t-1}^0(\omega) \in \mathcal{P}_{t-1}(\omega)$ and $E^{Q_{t-1}^0(\omega)}[\lambda_{t-1}^0(\omega)X_t(\omega, \cdot) + (1 - \lambda_{t-1}^0(\omega))Y_t(\omega, \cdot)] = \tilde{S}_{t-1}^B(\omega)$. Set $(Q_{t-1}, \lambda_{t-1}, \hat{P}_{t-1}) = (Q_{t-1}^0, \lambda_{t-1}^0, \hat{P}_{t-1}^0)$ on $C \setminus \hat{N}$, and $Q_{t-1} = \hat{P}_{t-1} = P_{t-1}$, $\lambda_{t-1} = 1/2$ on $C \cap \hat{N}$.

Then $Q_{t-1}(\cdot)$, $\lambda_{t-1}(\cdot)$, $\hat{P}_{t-1}(\cdot)$ constructed above are universally measurable. In all cases, define $\tilde{S}_t := (1 - \lambda_{t-1})X_t + \lambda_{t-1}Y_t$. Then \tilde{S}_t is obviously universally measurable, and we have $E^{Q_{t-1}(\omega)}[\tilde{S}_t(\omega, \cdot)] = \tilde{S}_{t-1}(\omega)$ Q' -a.s.. Define $Q := Q' \otimes Q_{t-1}$, $\hat{P} = \hat{P}' \otimes \hat{P}_{t-1}$, we then have $P \ll Q \ll \hat{P} \in \mathcal{P}$ (notice that $P = P|_{\Omega_{t-1}} \otimes P_{t-1} = P|_{\Omega_{t-1}} \otimes P_{t-1}^B$) and $E^Q[\tilde{S}_t | \mathcal{F}_{t-1}] = \tilde{S}_{t-1}$. That is, \tilde{S} is a generalized martingale since we do not necessarily have the integrability of \tilde{S}_t under Q . But by [55, Proposition 5.3.2] and [52, Theorem 1.1], we can replace Q by an equivalent probability measure under which \tilde{S} is a true martingale.

Step 3. We will now verify that the extended weight process $(\lambda_r)_{r=1, \dots, t-1}$ and the corresponding stopping times satisfies the property described in (ii'). We shall denote those stopping times for market \mathbf{M} by $\tau_n^0, \sigma_n^0, \tau_n^1, \sigma_n^1$. Notice that they differ

from their counterparts $\tilde{\tau}_n^0, \tilde{\sigma}_n^0, \tilde{\tau}_n^1, \tilde{\sigma}_n^1$ for market \mathbf{M}' only possibly in the last trading cycle. We check only the properties related to τ_n^0, σ_n^0 since our extension in step 2 is symmetric. In this step, to keep notation simple, we treat \mathcal{F}_r -measurable functions as defined on Ω for each r , i.e. if f is \mathcal{F}_r -measurable and $\omega|_{\Omega_r}$ is the first r components of $\omega \in \Omega$, then we write $f(\omega)$ to mean $f(\omega|_{\Omega_r})$.

Let $\omega \in N_{\mathbf{M}_1}^c \cap \{\tau_n^0 < \infty\}$. Also assume ω belongs to the \mathcal{P}' -q.s. set where (2.25) hold (the \mathcal{P}' -q.s. set is also \mathcal{P} -q.s.).

Case 1. $\tau_n^0(\omega) < \sigma_n^0(\omega) \leq t - 2$. In this case, $\tau_n^0(\omega) = \tilde{\tau}_n^0(\omega)$, $\sigma_n^0(\omega) = \tilde{\sigma}_n^0(\omega)$, and we have $X_{\tau_n^0(\omega)}(\omega) = Y_{\tau_n^0(\omega)}(\omega) = X_r(\omega)$ for $r \in [\tau_n^0(\omega), \sigma_n^0(\omega) \wedge t] = [\tilde{\tau}_n^0(\omega), \tilde{\sigma}_n^0(\omega) \wedge (t - 1)]$ by the induction hypothesis.

Case 2. $\tau_n^0(\omega) \leq T - 2$, $\sigma_n^0(\omega) = t - 1$. In this case, $\tau_n^0(\omega) = \tilde{\tau}_n^0(\omega)$ and $\tilde{\sigma}_n^0(\omega) = \infty$. Again, induction hypothesis gives $X_{\tau_n^0(\omega)}(\omega) = Y_{\tau_n^0(\omega)}(\omega) = X_r(\omega)$ for $r \in [\tau_n^0(\omega), \sigma_n^0(\omega) \wedge t] = [\tilde{\tau}_n^0(\omega), t - 1] = [\tilde{\tau}_n^0(\omega), \tilde{\sigma}_n^0(\omega) \wedge (t - 1)]$.

Case 3. $\tau_n^0(\omega) \leq t - 2$, $\sigma_n^0(\omega) = \infty$. In this case, $\tau_n^0(\omega) = \tilde{\tau}_n^0(\omega)$, $\tilde{\sigma}_n^0(\omega) = \infty$ and $[\tau_n^0(\omega), \sigma_n^0(\omega) \wedge t] = [\tilde{\tau}_n^0(\omega), t]$. Induction hypothesis implies $X_{\tau_n^0(\omega)}(\omega) = Y_{\tau_n^0(\omega)}(\omega) = X_r(\omega)$ for $r \in [\tilde{\tau}_n^0(\omega), t - 1]$. It remains to check $X_t(\omega) = X_{t-1}(\omega)$ for \mathcal{P} -q.s. such ω . In terms of λ process, case 3 corresponds to $\lambda_{t-2}(\omega) = \lambda_{t-1}(\omega) = 0$. $\lambda_{t-2}(\omega) = 0$ implies $\tilde{S}_{t-1}(\omega) = X_{t-1}(\omega)$. Based on our construction, $\lambda_{t-1}(\omega) = 0$ only when $\Lambda_F(\omega) = \{0\} \neq \Lambda_G(\omega)$. Since any $P \in \mathcal{P}$ has the decomposition $P = P|_{\Omega_{t-1}} \otimes P_{t-1}$, we must have $X_t(\omega) = \tilde{S}_{t-1}(\omega) = X_{t-1}(\omega)$ for \mathcal{P} -q.s. ω that fall into case 3.

Case 4. $\tau_n^0(\omega) = t - 1$, $\sigma_n^0(\omega) = \infty$. In this case, $\tilde{\tau}_n^0(\omega) = \tilde{\sigma}_n^0(\omega) = \infty$. We need to check $X_{t-1}(\omega) = Y_{t-1}(\omega) = X_t(\omega)$ for \mathcal{P} -q.s. such ω . In terms of λ process, case 4 corresponds to $\lambda_{t-2}(\omega) > 0 = \lambda_{t-1}(\omega)$. If $X_{t-1}(\omega) \neq Y_{t-1}(\omega)$, then $\lambda_{t-2}(\omega) > 0$ implies $\tilde{S}_{t-1}(\omega) > X_{t-1}(\omega)$ in which case $\Lambda_F(\omega) \cap \mathbb{R}_{<0} \neq \emptyset$. But based on our construction, $\lambda_{t-1}(\omega) = 0$ only if $\Lambda_F(\omega) = \{0\}$. So we must have $X_{t-1}(\omega) = Y_{t-1}(\omega) = \tilde{S}_{t-1}(\omega)$.

Similar to case 3, $X_t(\omega) = \tilde{S}_{t-1}(\omega) = X_{t-1}(\omega)$ for \mathcal{P} -q.s. ω that fall into case 4 follows from $\Lambda_F(\omega) = \{0\}$.

Statements about σ_n^1, τ_n^1 can be verified by a symmetric argument. We therefore have proved that (i) implies (ii') for the recursively defined markets $[X, Y]$ with t periods.

Finally, we note that (ii') clearly implies (ii).

(ii) \Rightarrow (i): Let $f \in A_T$ be such that $f \geq 0$ \mathcal{P} -q.s.. To show $f = 0$ \mathcal{P} -q.s., we suppose on the contrary $\exists P \in \mathcal{P}$ such that $P(\|f\| > 0) > 0$ and try to derive a contradiction. Write $f = \phi_{T+1}$ for some $\phi \in \mathcal{H}$. Let (Q, \tilde{S}) be the CPS given by (ii). It is easy to see that $(1, \tilde{S})$ is a Q -martingale which lies in $K^* \cap \mathbb{R}_{>0}^2$ where for each t and ω , $K_t^*(\omega) := \{y \in \mathbb{R}^d : \langle x, y \rangle \geq 0 \forall x \in K_t(\omega)\}$ is the dual cone of the solvency cone $K_t(\omega)$. According to Lemma A.3.1, we can pick $Q' \sim Q$ such that $\Delta\phi_{t+1}$ are Q' -integrable for all $t = 0, \dots, T$. Then by a slight modification of [55, Lemma 3.2.4] (simply replace $\text{ri } G^*$ with $K^* \cap \mathbb{R}_{>0}^2$), there exists a bounded Q' -martingale $Z = (Z^0, Z^1) \in K^* \cap \mathbb{R}_{>0}^2$.

Now, let $\langle \cdot, \cdot \rangle$ denote the usual inner product. On one hand,

$$E^{Q'}[\langle Z_T, f \rangle] = \sum_{t=0}^T E^{Q'}[\langle Z_t, \Delta\phi_{t+1} \rangle] = \sum_{t=0}^T E^{Q'}[\langle Z_t, \Delta\phi_{t+1} \rangle] \leq 0$$

by the martingale property of Z under Q' and that $Z_t \in K_t^*$, $\Delta\phi_{t+1} \in -K_t$. On the other hand, $Q' \sim Q \gg P$ implies $Q'(\|f\| > 0) > 0$. Together with $f \geq 0$ and $Z_T \in \mathbb{R}_{>0}^2$ we get the contradictory inequality:

$$E^{Q'}[\langle Z_T, f \rangle] > 0.$$

□

Proof of Theorem 2.3.7. $\text{NA}^r(\mathcal{P})$ holds implies that $\text{NA}(\mathcal{P})$ holds for a less frictional market $[\underline{S}', \bar{S}'] \subseteq \text{ri}[\underline{S}, \bar{S}]$ where $\underline{S}_t, \bar{S}_t$ are also continuous for each t . Then apply

Theorem 2.3.6 to the $[\underline{S}', \overline{S}']$ -market. The reverse implication is similar to that of Theorem 2.3.6 with $[\underline{S}', \overline{S}']$ -market as the $NA(\mathcal{P})$ candidate. \square

2.3.4 Existence of an optimal superhedging strategy

In this subsection, we give an existence result of an optimal superhedging strategy when the market satisfies $NA^r(\mathcal{P})$. The proof is based on the closedness of A_T under \mathcal{P} -q.s. convergence. We begin with a quasi-sure version of [71, Lemma 2.6] the proof of which is almost identical to its classical version and is included in Appendix A.4 for the readers' convenience.

For a random set $A \subseteq \mathbb{R}^2$, write $L^0(A; \mathcal{F}_t)$ for the set of \mathcal{F}_t -measurable random vectors taking valued in A . Also let $K_t^0 := K_t \cap (-K_t)$ be the linear space of portfolios that can be converted to zero and vice versa.

Lemma 2.3.17. *Suppose $NA^r(\mathcal{P})$ holds. Let $\xi_t \in L^0(-K_t; \mathcal{F}_t), t = 0, \dots, T$ satisfy $\sum_{t=0}^T \xi_t = 0$ \mathcal{P} -q.s.. Then $\xi_t \in K_t^0$ for $t = 0, \dots, T$ \mathcal{P} -q.s..*

Proposition 2.3.18. *If $NA^r(\mathcal{P})$ holds, then A_T is closed under \mathcal{P} -q.s. convergence.*

Proof. Let $\{W^n\}$ be a sequence in A_T which converges to a random variable W \mathcal{P} -q.s.. We want to show $W = \sum_{t=0}^T \xi_t$ \mathcal{P} -q.s. for some process ξ satisfying $\xi_t \in L^0(-K_t; \mathcal{F}_t), t = 0, \dots, T$. We mimic the proof of Theorem 2.2 in [25] and do induction on the number of periods in the market. In this proof, we do not assume Ω_0 is a singleton.

When the market has zero period, each $W^n \in -K_0$ \mathcal{P} -q.s.. Since K_0 is closed-valued, $W \in -K_0$ \mathcal{P} -q.s., and we are done. Suppose the claim is true for any market with dates $\{1, 2, \dots, T\}$, we now deduce the case with dates $\{0, 1, \dots, T\}$. Note that $NA^r(\mathcal{P})$ for market with dates $\{0, 1, \dots, T\}$ clearly implies $NA^r(\mathcal{P})$ for market with dates $\{1, \dots, T\}$. We can write $W^n = \sum_{t=0}^T \xi_t^n - (g^n, 0)$ \mathcal{P} -q.s. with

$\xi_t^n \in L^0(-K_t; \mathcal{F}_t)$, $\xi_t^{n,0} = -(\xi_t^{n,1})^+ \bar{S}_t + (\xi_t^{n,1})^- \underline{S}_t$, $t = 0, \dots, T$, and $g^n \in L^0(\mathbb{R}_{\geq 0}; \mathcal{F}_T)$.

Here g^n represents the total amount of cash thrown away up to time T .

Consider the \mathcal{F}_0 -measurable set $E := \{\liminf_n |\xi_0^{n,1}| < \infty\}$. We can find \mathcal{F}_0 -measurable random indices n_k such that $\xi_0^{n_k} = (-(\xi_0^{n_k,1})^+ \bar{S}_0 + (\xi_0^{n_k,1})^- \underline{S}_0, \xi_0^{n_k,1})$ converges pointwise to a finite \mathcal{F}_0 -measurable random vector $\xi_0 \in -K_0$.

On E , we have

$$\sum_{t=1}^T \xi_t^{n_k} - (g^{n_k}, 0) \rightarrow W - \xi_0 \quad \mathcal{P}\text{-q.s.}$$

By induction hypothesis, $\exists \xi_t \in L^0(-K_t; \mathcal{F}_t)$, $t = 1, \dots, T$ such that

$$(2.26) \quad W - \xi_0 = \sum_{t=1}^T \xi_t \quad \mathcal{P}\text{-q.s. on } E.$$

If $E^c = \{\liminf_n |\xi_0^{n,1}| = \infty\}$ is \mathcal{P} -polar, then we are done. Suppose E^c is not \mathcal{P} -polar, let $G_0^n := \frac{\xi_0^n}{1 + |\xi_0^{n,1}|}$. Since $|G_0^{n,1}| \leq 1$ and $|G_0^{n,0}| = \frac{|-(\xi_0^{n,1})^+ \bar{S}_0 + (\xi_0^{n,1})^- \underline{S}_0|}{1 + |\xi_0^{n,1}|} \leq \bar{S}_0$, there exists \mathcal{F}_0 -measurable random indices n_k such that $G_0^{n_k}$ converges pointwise to an \mathcal{F}_0 -measurable random vector $G_0 = (G_0^0, G_0^1)$ with $|G_0^1| = 1$ on E^c . $G_0 \in -K_0$ since $-K_0$ is a (random) closed cone. Divide by $1 + |\xi_0^{n_k,1}|$ on both sides of $\sum_{t=1}^T \xi_t^{n_k} - (g^{n_k}, 0) = W^{n_k} - \xi_0^{n_k}$ and take limit as $k \rightarrow \infty$, we get

$$\sum_{t=1}^T \frac{\xi_t^{n_k}}{1 + |\xi_0^{n_k,1}|} - \frac{(g^{n_k}, 0)}{1 + |\xi_0^{n_k,1}|} \rightarrow -G_0 \quad \mathcal{P}\text{-q.s. on } E^c.$$

Clearly, $\frac{\xi_t^{n_k}}{1 + |\xi_0^{n_k,1}|} \in L^0(-K_t; \mathcal{F}_t)$, $t = 1, \dots, T-1$ and $\frac{\xi_T^{n_k} - (g^{n_k}, 0)}{1 + |\xi_0^{n_k,1}|} \in L^0(-K_T; \mathcal{F}_T)$. So we can apply induction hypothesis again to obtain $\tilde{\xi}_t \in L^0(-K_t; \mathcal{F}_t)$, $t = 1, \dots, T$ such that

$$(2.27) \quad G_0 + \sum_{t=1}^T \tilde{\xi}_t = 0 \quad \mathcal{P}\text{-q.s. on } E^c.$$

Observe that K_t^0 is either a single point (the origin) if $\underline{S}_t < \bar{S}_t$, or a line passing through the origin if $\underline{S}_t = \bar{S}_t$. By Lemma 2.3.17, we have that on E^c , $G_0 \in K_0^0$ and

$\tilde{\xi}_t \in K_t^0$ for all $t = 1, \dots, T$ \mathcal{P} -q.s.. Since $G_0 \neq 0$ on E^c by construction, we must have $K_0^0 \neq \{0\}$ \mathcal{P} -q.s. on E^c . For each $t = 0, \dots, T$, write $\bar{S}_t = \underline{S}_t := S_t$ whenever $K_t^0 \neq \{0\}$. On E^c , we have \mathcal{P} -q.s., $G_0^0 + G_0^1 S_0 = 0$, and for $t = 1, \dots, T$, either $\tilde{\xi}_t \in K_t^0 = \{0\}$, or $\tilde{\xi}_t^0 + \tilde{\xi}_t^1 S_t = 0$ whenever $K_t^0 \neq \{0\}$. Define

$$\eta_0^n := 1_E \xi_0^n,$$

and for $t = 1, \dots, T$,

$$\eta_t^{n,1} := \xi_t^{n,1} - 1_{E^c} \frac{\xi_0^{n,1}}{G_0^1} \tilde{\xi}_t^1, \quad \eta_t^{n,0} := -(\eta_t^{n,1})^+ \bar{S}_t + (\eta_t^{n,1})^- \underline{S}_t.$$

Then $\eta_t^n \in L^0(-K_t; \mathcal{F}_t)$, and

$$\sum_{t=0}^T \eta_t^{n,1} = \sum_{t=0}^T \xi_t^{n,1} - 1_{E^c} \frac{\xi_0^{n,1}}{G_0^1} \left(G_0^1 + \sum_{t=1}^T \tilde{\xi}_t^1 \right) = \sum_{t=0}^T \xi_t^{n,1} \quad \mathcal{P}\text{-q.s.},$$

where the last equality holds by (2.27). On E , we have

$$\sum_{t=0}^T \eta_t^{n,0} = \xi_0^{n,0} + \sum_{t=1}^T -(\xi_t^{n,1})^+ \bar{S}_t + (\xi_t^{n,1})^- \underline{S}_t = \xi_0^{n,0} + \sum_{t=1}^T \xi_t^{n,0} = \sum_{t=0}^T \xi_t^{n,0}.$$

On E^c , we have

$$\begin{aligned} \sum_{t=0}^T \eta_t^{n,0} &= \sum_{t=1}^T -(\eta_t^{n,1})^+ \bar{S}_t + (\eta_t^{n,1})^- \underline{S}_t \\ &\geq \sum_{t=1}^T -(\xi_t^{n,1})^+ \bar{S}_t + (\xi_t^{n,1})^- \underline{S}_t - \left(-\frac{\xi_0^{n,1}}{G_0^1} \tilde{\xi}_t^1 \right)^+ \bar{S}_t + \left(-\frac{\xi_0^{n,1}}{G_0^1} \tilde{\xi}_t^1 \right)^- \underline{S}_t \\ &= \sum_{t=1}^T \xi_t^{n,0} - \left(\frac{\xi_0^{n,1}}{G_0^1} \tilde{\xi}_t^1 \right)^- \bar{S}_t + \left(\frac{\xi_0^{n,1}}{G_0^1} \tilde{\xi}_t^1 \right)^+ \underline{S}_t \\ &= \sum_{t=1}^T \xi_t^{n,0} - \frac{\xi_0^{n,1}}{G_0^1} \sum_{t=1}^T \tilde{\xi}_t^0 \\ &= \sum_{t=1}^T \xi_t^{n,0} + \frac{\xi_0^{n,1} G_0^0}{G_0^1} \\ &= \sum_{t=1}^T \xi_t^{n,0} - \xi_0^{n,1} S_0 = \sum_{t=0}^T \xi_t^{n,0}. \end{aligned}$$

In the second line, we used that the function $x \mapsto -x^+ \bar{S}_t + x^- \underline{S}_t$ is concave. In the fourth line, we used $\tilde{\xi}_t = 0$ when $K_t^0 = \{0\}$, and $\bar{S}_t = \underline{S}_t = S_t$, $\tilde{\xi}_t^0 + \tilde{\xi}_t^1 S_t = 0$ when $K_t^0 \neq \{0\}$. In the fifth line, we used (2.27). In the last line, we used $G_0^0 + G_0^1 S_0 = 0$.

So we have constructed $\eta_t^n \in L^0(-K_t; \mathcal{F}_t)$ such that $\eta_0^n = 0$ on E^c and

$$\sum_{t=0}^T \eta_t^n - \left(g^n + \sum_{t=0}^T \eta_t^{n,0} - \sum_{t=0}^T \xi_t^{n,0}, 0 \right) = \sum_{t=0}^T \xi_t^n - (g^n, 0) \rightarrow W \quad \mathcal{P}\text{-q.s.}$$

Now we can apply the induction hypothesis to obtain $\xi_t \in L^0(-K_t; \mathcal{F}_t)$ such that $W = \sum_{t=1}^T \xi_t$ \mathcal{P} -q.s. on E^c . Combining the result on E^c with that on E (see (2.26)), we have proved that $W \in A_T$. \square

Given an \mathbb{R}^2 -valued random variable f which will be treated as a contingent claim, define its superhedging price

$$\pi(f) := \inf\{x \in \mathbb{R} : \exists \phi \in \mathcal{H} \text{ such that } (x, 0) + \phi_{T+1} \geq f \text{ } \mathcal{P}\text{-q.s.}\}.$$

Theorem 2.3.19. *Let $NA^r(\mathcal{P})$ hold and let f be an \mathbb{R}^2 -valued random variable.*

Then $\pi(f) > -\infty$, and there exists $\phi \in \mathcal{H}$ such that $(\pi(f), 0) + \phi_{T+1} \geq f$ \mathcal{P} -q.s..

Proof. If $\pi(f) = \infty$, there is nothing to prove. If $|\pi(f)| < \infty$, then $f - (\pi(f) + 1/n, 0) \in A_T \forall n \in \mathbb{N}$. By closedness of A_T , $f - (\pi(f), 0) \in A_T$, meaning $\exists \phi \in \mathcal{H}$ such that $(\pi(f), 0) + \phi_{T+1} = f$ \mathcal{P} -q.s.. We now show that $\pi(f) = -\infty$ violates our no-arbitrage assumption. Suppose $\pi(f) = -\infty$, then for each $n \in \mathbb{N}$, $\exists \phi^n \in \mathcal{H}$ such that $(-n, 0) + \sum_{t=0}^T \Delta \phi_{t+1}^n \geq f$ \mathcal{P} -q.s.. Fix $0 < \epsilon < 1/\bar{S}_0$ and define $\eta_0^n := 0$, $\Delta \eta_1^n := \Delta \phi_1^n + n(-\epsilon \bar{S}_0, \epsilon)$ and $\Delta \eta_t^n := \Delta \phi_t^n$ for $t = 2, \dots, T+1$. It is easy to see that $(-\epsilon \bar{S}_0, \epsilon) \in -K_0$ and hence $\Delta \eta_1^n \in -K_0$. So $\eta^n \in \mathcal{H} \forall n$. Moreover, $f \leq (-n, 0) + \sum_{t=0}^T \Delta \phi_{t+1}^n = \sum_{t=0}^T \Delta \eta_{t+1}^n - n(1 - \epsilon \bar{S}_0, \epsilon)$ \mathcal{P} -q.s.. So we get

$$\sum_{t=0}^T \Delta \eta_{t+1}^n \geq f + n(1 - \epsilon \bar{S}_0, \epsilon) \geq ((f^0 + n(1 - \epsilon \bar{S}_0)) \wedge 1, (f^1 + n\epsilon) \wedge 1) \quad \mathcal{P}\text{-q.s.}$$

It follows from the solidness of A_T that $((f^0 + n(1 - \epsilon \bar{S}_0)) \wedge 1, (f^1 + n\epsilon) \wedge 1) \in A_T$. Let $n \rightarrow \infty$. Proposition 2.3.18 then implies $(1, 1) \in A_T$, violating $\text{NA}(\mathcal{P})$, and hence $\text{NA}^r(\mathcal{P})$. \square

2.4 Further discussion

2.4.1 Challenges with multi-asset extension

When there are multiple dynamically traded assets with bid-ask spreads, the market can be modeled by a so-called bid-ask matrix $(\pi_t^{ij}) : \Omega_t \rightarrow M^{d \times d}(\mathbb{R}_{>0})$ whose ij -th entry represents the number of physical units of the i -th asset needed to exchange for one unit of the j -th asset. In this set-up, and assuming the market has efficient and bounded friction, Bouchard and Nutz proved in a recent paper [24] a version of the FTAP with no arbitrage of the second kind, denoted by $\text{NA}_2(\mathcal{P})$. This no-arbitrage notion, however, is quite strong. It means the market is already in a good form for martingale extension, so that the backward recursion is avoided. A simple one-period market $\underline{S}_0 = 1, \bar{S}_0 = 3, \underline{S}_1 = 2, \bar{S}_1 = 4$ which satisfies $\text{NA}(\mathcal{P})$, thus reasonable in our opinion, fails $\text{NA}_2(\mathcal{P})$.

Rather than assuming $\text{NA}_2(\mathcal{P})$, it is desirable to work with a weaker no-arbitrage notion, e.g. strict no-arbitrage $\text{NA}^s(\mathcal{P})$, and obtain $\text{NA}_2(\mathcal{P})$ through backward recursion. There seem to be three possible ways to define a modified market.

- (i) Do backward recursion on the solvency cones via

$$\tilde{K}_t = \overline{K_t + \Lambda_t} \quad \text{where} \quad \Lambda_t(\omega) = \bigcap_{\omega' \in \text{supp} \mathcal{P}_t(\omega)} \tilde{K}_{t+1}(\omega, \omega').$$

- (ii) Do recursion on the dual cones via

$$\tilde{K}_t^* = K_t^* \cap \Lambda_t^* \quad \text{where} \quad \Lambda_t^*(\omega) = \overline{\text{conv} \tilde{K}_{t+1}^*(\omega, \text{supp} \mathcal{P}_t(\omega))}.$$

Here “*conv*” stands for convex hull.

(iii) Do recursion on the bid-ask matrix via

$$\tilde{\pi}_t^{ij} = \gamma_t \wedge \pi_t^{ij} \vee \frac{1}{\pi_t^{ji}} \quad \text{where} \quad \gamma_t^{ij}(\omega) := \sup \supp_{\mathcal{P}_t(\omega)} \tilde{\pi}_{t+1}^{ij}(\omega, \cdot),$$

and then generate the corresponding solvency cone \tilde{K}_t .

The issue with methods (i) and (ii) is that the intersection of continuous set-valued maps need not be continuous, not to mention that we generally have an uncountable intersection in (i). Since measurability is preserved under countable intersection of closed-valued maps, it may appear that one can work with measurable rather than continuous set-valued maps in (ii). In fact, (ii) is similar to the recursion used in [66]. However, the composition of two measurable set-valued map may fail to be measurable unless the outer map \tilde{K}_{t+1}^* is of Carathéodory type, i.e. measurable in ω and continuous in ω' .

Method (iii) shares the most similarity with Section 2.3. The nice thing with doing recursion on the generators instead of the cones is that continuity is now preserved. The modified market defined in this way does preserve the no-arbitrage property of the original market, and is in some sense, least favorable, but it is not in a very good form for martingale extension. For martingale extension, we want $\text{NA}_2(\mathcal{P})$ to hold locally, which would be the case if the solvency cone generated by the worst exchange rates γ_t^{ij} is equal to the worst solvency cone $\bigcap_{\omega' \in \supp \mathcal{P}_t(\omega)} \tilde{K}_{t+1}(\omega, \omega')$. This is true is dimension two, but problematic in higher dimensions.

2.4.2 Challenges with superhedging duality

In the quasi-sure framework, superhedging theorem without friction is proved in [25] using a dynamic programming approach where the martingale measures serve as the control. Compared with the frictionless case where everything is measured in

cash, in the frictional case, we need to work with multi-dimensional positions measured in physical units. Then at each time t in the dynamic programming procedure, instead of having a single point (the smallest initial capital) from which one can superhedge the value function at time $t + 1$, we now have a set G of minimal points. This means the dimension of the problem increases: there is an additional minimization problem over G ; we pick the element in G which has the smallest superhedging price in the first t -period market.

We have an “inf sup” problem where we minimize over random variables in G and maximize over consistent price systems. To make the inductive step work and use the one-period result, we need the exchangeability of “inf” and “sup”, i.e. a minimax theorem of some kind. But neither G (equipped with the topology of pointwise convergence) nor the set of consistent price systems are compact.

CHAPTER III

Minimizing the Probability of Lifetime Ruin Under Ambiguity Aversion

3.1 Introduction

The problem of how individuals should invest their wealth in a risky financial market to minimize the probability that they outlive their wealth, also known as the *probability of lifetime ruin* (this term was coined by [62]), was analyzed by Young [82]. We mention that Jacka in an earlier work [47] considered a finite-fuel problem of very similar form. Subsequent variants of Young's work include but not limited to adding borrowing constraints [14], assuming consumption is ratcheted [15], allowing stochastic consumption [16] and stochastic volatility [7]. In all previous works, there is a fixed risky asset model; that is, the investor is certain about the evolution and distribution of the risky asset price. This is, however, not very realistic. There may be good estimates of the price volatility, but drift estimation, as Rogers points out in [64, Section 4.2], is almost impossible; it would require centuries of data to obtain a reliable estimate. Therefore, it is desirable to have a robust investment strategy that can perform well against drift misspecification. For a good introduction of robust decision making theory, see [40].

Although drift estimation is difficult, one would still like to make use of the

available data. A natural approach is to extract from the available data a reference model, and penalize other models based on their deviation from the reference model. How hard to penalize depends on how averse the agent is to *ambiguity*, also called *model uncertainty* or *Knightian uncertainty*. Early works incorporating ambiguity aversion into optimization (e.g. [60], [41]) are mostly done via a formal analysis of the corresponding Hamilton-Jacobi-Bellman (HJB) equation. Among those that provide more mathematical rigor, we mention a few that use different approaches. Jaimungal solves a finite horizon irreversible investment problem [49], and a hybrid model of default problem with Sigloch [50] using stochastic control. They work with a scaled entropic penalty in order to get explicit solutions and rely on direct verification. Bordigoni et al. [23] analyze a finite horizon utility maximization problem also by control method, but provide a backward stochastic differential equation (BSDE) characterization instead of an HJB characterization. Their results are generalized to an infinite horizon setting by Hu and Schweizer [44]. Schied [72] and Hernández-Hernández and Schied [42] treat robust utility maximization problems using duality or a combination of duality and control.

In this chapter, we provide a complete and rigorous analysis of the robust lifetime ruin problem

$$\inf_{\pi} \sup_{\mathbb{Q}} \left\{ \mathbb{Q}(\tau_b < \tau_d) - \frac{1}{\varepsilon} h^d(\mathbb{Q}|\mathbb{P}) \right\}$$

using stochastic control, where τ_b and τ_d are the ruin time and death time, respectively, h^d is a variant of the entropic penalty function which only measures entropy up to the death time, ε specifies the penalization strength, π runs through a set of investment strategies and \mathbb{Q} runs through a set of possible models representing drift uncertainty. When the hazard rate is zero, we obtain explicit formulas. In the general case, we characterize the value function as the unique classical solution of

an associated HJB equation satisfying two boundary conditions, and give feedback forms for the optimal investment and drift distortion. In contrast to the non-robust case or robust utility maximization problem, we show that the value function loses convexity for a class of parameters, which suggests that the Isaacs condition may fail. Same as the non-robust case, we also show that the optimally controlled wealth process never reaches the so-called “safe level”. So the goal is to stay away from the ruin level and to “win” the game by dying. When the hazard rate is zero, the inaccessibility of the safe level is pointed out in [26]. Goal reaching problems without a deadline goes back to the work of Pestien and Sudderth [63] (also see [5]). The optimal strategy there is to maximize the ratio of drift to volatility squared. We shall see that in terms of the optimal investment strategy, the robustness is only non-trivial when death is added.

Our work extends the discussion in [82] to the robust case. Unlike [49] and [50] where a scaled entropic penalty leads to explicit solutions, our random horizon robust problem, even in the simple Black-Scholes framework, fails to have an explicit solution in general, whether the penalty is scaled or not. Moreover, due to degeneracy and the control space being unbounded, the classical nonlinear elliptic theory by Krylov [58] cannot be applied directly. So we have to resort to the theory of viscosity solutions and then upgrade regularity by bootstrapping. Our work differs from [23], [44], [72], [42] in the methodology. The BSDE characterizations in [23] and [44] only focus on the inner \mathbb{Q} -maximization problem and do not describe the optimal investment strategy or the saddle point. The duality approach of [72] requires the infimum and supremum to be exchangeable, which does not hold in our case with certain choice of parameters. The classical duality $\log \mathbb{E}[e^X] = \sup_{\mathbb{Q} \in \mathcal{Q}_{abs}} \{\mathbb{E}^{\mathbb{Q}}[X] - h(\mathbb{Q}|\mathbb{P})\}$ between free energy and entropic penalty may look useful at a first glance, but the

uncertainty set \mathcal{Q}_{abs} does not preserve the independence between asset price and mortality, and does not leave room for varied confidence levels regarding different model components.¹ In addition, we are not using the exact entropic function h , but its variant h^d . Due to time-inconsistency issue, we do not consider uncertainty in hazard rate. It could be an interesting extension to have uncertain Poisson jump rate (see e.g. [59], [61], [27], [21]) and to allow varied positive levels of ambiguity aversion (see e.g. [81], [49]).

For the construction of a viscosity solution to the HJB equation, we use a “comparison + Perron’s method” approach described in [29] instead of the usual route of “dynamic programming principle (DPP) + value function is a viscosity solution + comparison”. The reason is that robust optimization problems resemble stochastic differential games in which nature can be regarded as the second player, and the DPP for games is generally complicated because of measurability issues. One either has to use the Elliott-Kalton formulation where one player uses controls and the other player uses “strategies”, i.e. maps defined on a set of controls satisfying nonanticipativity (see e.g. [38], [37], [12]), or restrict oneself to strategies of simple form, for example, to what Sîrbu [74] calls elementary strategies. Both ways to get around the measurability issues are not ideal for us. In particular, it is a bit unnatural for us to use the Elliott-Kalton formulation and assume nature is a strategic player against us, because nature has no payoff and is disinterested. It turns out that the classical Perron’s method yields a much simpler and more elegant construction. The only drawback is that regularity now becomes very important, otherwise the constructed solution cannot be related to the value function. Fortunately, we are able to upgrade regularity and carry out a verification theorem. The approach outlined here was first

¹ \mathcal{Q}_{abs} denotes the set of measures that are absolutely continuous with respect to \mathbb{P} and have finite entropy.

used by Janeček and Sîrbu [51] in a pure stochastic control problem.

Convexity is usually key to upgrading regularity. One challenge introduced by robustness, as we have pointed out, is the loss of convexity of the value function for a class of parameters. In fact, even for non-robust lifetime ruin problems, *a priori* convexity of the value function is not clear. For example, [16] obtains convexity for a lifetime ruin problem with stochastic consumption by going to a controller-and-stopper problem whose convex dual is related to the original problem through a dimension reduction. We overcome this challenge by working with an equivalent convex problem obtained through the Cole-Hopf transformation. Once we have convexity, it is easy to upgrade to C^1 -regularity using convex analysis and the theory of viscosity solutions. We further upgrade to C^2 -regularity by analyzing a Poisson equation, where we borrow some techniques from [51] and [73]. One may try to prove C^2 -regularity by the regularization method used in [83] and [36], but such an approach requires us to prove the existence of a positive lower bound on π that is independent of the regularization on compact intervals away from the safe level, which we find to be difficult to establish.

The rest of this chapter is organized as follows. In Section 4.2, we set up the problem, derive the HJB equation and feedback forms heuristically, and state the main results. Section 3.3 provides an explicit solution when the hazard rate is zero, which is not only interesting for its own sake, but serves as a useful upper bound in the analysis of the general case. Sections 3.4 and 3.5 are devoted to establishing the existence of a classical solution to the HJB equation, with Section 3.4 focusing on Perron's construction of a viscosity solution, and Section 3.5 on regularity. In Section 3.6, we give a verification theorem and the proof of our main results. In order to prove verification theorem, we also show the boundedness and Lipschitz

continuity of the optimal investment strategy. Sections 3.7 collects some additional properties of the optimal investment strategy and the value function. Sections 3.8 and 3.9 provides formulas for small ε -expansion and numerical examples.

3.2 Problem formulation

Let Ω^M be the space of continuous functions $\omega : [0, \infty) \rightarrow \mathbb{R}$, equipped with the topology of uniform convergence on compact subintervals of $[0, \infty)$. Let \mathcal{F}^M be the Borel sigma-algebra on Ω^M and \mathbb{P}^M be the Wiener measure on $(\Omega^M, \mathcal{F}^M)$. The coordinate map $B_t(\omega) := \omega(t)$ is a standard Brownian motion in this space. Here \mathbb{P}^M serves as a reference measure which reflects an individual's belief about the market. Let $N = (N_t)_{t \geq 0}$ be a Poisson process with rate λ defined on another probability space $(\Omega^d, \mathcal{F}^d, \mathbb{P}^d)$. Let τ_d be the first time that the Poisson process jumps, modeling the death time of the individual. τ_d is an exponential random variable with parameter λ which is known as the *hazard rate* in this context. Define

$$(\Omega, \mathcal{F}, \mathbb{P}) := (\Omega^M \times \Omega^d, \mathcal{F}^M \otimes \mathcal{F}^d, \mathbb{P}^M \times \mathbb{P}^d).$$

B and N are independent on this space, and remain a Brownian motion and a Poisson process, respectively. Let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be the (raw) filtration generated by the Brownian motion B and $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ be the filtration generated by B and the process $1_{\{\tau_d \leq t\}}$. Assume both \mathbb{F} and \mathbb{G} have been made right continuous. However, we do not complete the filtrations because later on, we would like to include measures that are only locally equivalent to \mathbb{P} as part of our consideration.²

The individual invests in a financial market consists of a risk-free bank account with interest rate $r > 0$ and a risky asset whose price S_t follows a geometric Brownian

²By locally equivalent, we mean equivalent on \mathcal{G}_t for all $t \geq 0$. Although the filtrations in our setup is not complete, stochastic integral can still be defined and has all the usual properties. In particular, Itô's lemma is still valid. See, for example, chapter 1 of [48].

motion:

$$dS_t = \mu S_t + \sigma S_t dB_t, \quad S_0 = S > 0$$

where $\mu > r$ and $\sigma > 0$. Let π_t be the amount that the individual invests in the risky asset at time t . Apart from investment, the individual also consumes at a constant rate $c > 0$ of her current wealth w .³ Her wealth W_t evolves according to the stochastic differential equation (SDE):

$$dW_t = [rW_t + (\mu - r)\pi_t - c]dt + \sigma\pi_t dB_t, \quad W_0 = w.$$

An investment strategy π is admissible if it is \mathbb{F} -progressively measurable and almost surely bounded (uniformly in time).⁴ Denote by \mathcal{A} the set of all admissible strategies.

Let $\tau_b := \inf\{t \geq 0 : W_t \leq b\}$ be the first time the individual's wealth falls to or below a specified ruin level b . The individual aims at minimizing the probability that ruin happens before death, i.e. $\tau_b < \tau_d$, in a robust sense. More precisely, she suspects that the drift of the risky asset may be misspecified. So instead of optimizing under the reference measure \mathbb{P} , she considers a set \mathcal{Q} of candidate measures that are locally equivalent to \mathbb{P} , and penalizes their deviation from \mathbb{P} . Here we assume the individual is only robust against the market model, but not the death time model, nor the independence between them. So elements in \mathcal{Q} should be of the form $\mathbb{Q}^M \times \mathbb{P}^d$ so that τ_d remains an $exp(\lambda)$ random variable under all candidate measures. Let $h(\mathbb{Q}|\mathbb{P}) := \mathbb{E}^{\mathbb{Q}}[\log \frac{d\mathbb{Q}}{d\mathbb{P}}]$ be the relative entropic function. Denote by \mathbb{Q}_t the restriction of a measure \mathbb{Q} to \mathcal{G}_t . We penalize the deviation from \mathbb{P} using a variant of h :

$$h^d(\mathbb{Q}|\mathbb{P}) := h(\mathbb{Q}_{\tau_d}|\mathbb{P}_{\tau_d})$$

³To simplify the discussion, we only work with constant consumption rate. But the main techniques can be applied to proportional consumption rate, and more generally, to the case when the consumption rate is a non-negative, Lipschitz continuous function of wealth.

⁴Almost sure boundedness can be relaxed as long as the best drift distortion in response to each π defines an admissible measure $\mathbb{Q} \in \mathcal{Q}$ where \mathcal{Q} is the model uncertainty set to be introduced.

which only measures the relative entropy on \mathcal{G}_{τ_d} ; that is, the individual does not care about drift uncertainty after death. She faces the following robust optimization problem:

$$(3.1) \quad \psi(w; \varepsilon) = \inf_{\pi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{Q}_w(\tau_b < \tau_d) - \frac{1}{\varepsilon} h^d(\mathbb{Q}|\mathbb{P}) \right\},$$

where the subscript w represents conditioning on the event $W_0 = w$. The parameter ε measures the individual's level of ambiguity aversion or preference for robustness. $\varepsilon \downarrow 0$ corresponds to the classical non-robust case since all measures other than \mathbb{P} would give a very negative value, thus not optimal for the inner maximization problem. A larger ε means the individual is more ambiguity averse, has less faith in the reference model and will consider larger drift distortion. $\varepsilon \rightarrow \infty$ corresponds to the worst-case approach, i.e. the individual has equal belief in all candidate measures and optimize again the worst-case scenario.

We now give the precise definition of the set \mathcal{Q} of candidate measures. A probability measure $\mathbb{Q} \in \mathcal{Q}$ if

$$(3.2) \quad \frac{d\mathbb{Q}_t}{d\mathbb{P}_t} = \exp \left(-\frac{1}{2} \int_0^t \boldsymbol{\theta}_s^2 ds + \int_0^t \boldsymbol{\theta}_s dB_s \right), \quad t \geq 0$$

for some \mathbb{F} -progressively measurable process $\boldsymbol{\theta}$ satisfying $\mathbb{E}[e^{\frac{1}{2} \int_0^t \boldsymbol{\theta}_s^2 ds}] < \infty$ for all $t \geq 0$, and $\mathbb{E}^{\mathbb{Q}}[\int_0^\infty e^{-\lambda s} \boldsymbol{\theta}_s^2 ds] < \infty$. Conversely, given any \mathbb{F} -progressively measurable process $\boldsymbol{\theta}$ satisfying $\mathbb{E}[e^{\frac{1}{2} \int_0^t \boldsymbol{\theta}_s^2 ds}] < \infty$ for all $t \geq 0$, we can define a consistent family of measures $\mathbb{Q}_t \sim \mathbb{P}_t$ on (Ω, \mathcal{G}_t) by (3.2). By [77, Lemma 4.2] (also see [45, Proposition 1]), there exists a probability measure \mathbb{Q} on (Ω, \mathcal{F}) such that $\mathbb{Q}|_{\mathcal{G}_t} = \mathbb{Q}_t$ for all $t \geq 0$.⁵ Throughout this paper, we will use boldface greks $\boldsymbol{\pi}, \boldsymbol{\theta}$ to denote controls (as stochastic processes) and plain greks π, θ to denote the values that the controls can take. Since τ_d is independent of \mathbb{F} , the distribution of τ_d is invariant under such

⁵The existence of such a measure is not guaranteed if the filtration has been completed w.r.t. \mathbb{P} .

change of measure. Under \mathbb{Q} , S_t has drift $\mu + \sigma\theta_t$ and W_t has dynamics:

$$(3.3) \quad dW_t = [rW_t + (\mu + \sigma\theta_t - r)\pi_t - c]dt + \sigma\pi_t dB_t^{\mathbb{Q}}$$

where $B^{\mathbb{Q}}$ is a \mathbb{Q} -Brownian motion independent of τ_d .

Let $\mathbb{Q} \in \mathcal{Q}$. We have

$$\begin{aligned} h^d(\mathbb{Q}|\mathbb{P}) &= \mathbb{E}^{\mathbb{Q}} \left[-\frac{1}{2} \int_0^{\tau_d} \theta_s^2 ds + \int_0^{\tau_d} \theta_s dB_s \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[-\frac{1}{2} \int_0^{\tau_d} \theta_s^2 ds + \int_0^{\tau_d} \theta_s (dB_s^{\mathbb{Q}} + \theta_s ds) \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{2} \int_0^{\tau_d} \theta_s^2 ds \right] = \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{2} \int_0^{\infty} e^{-\lambda s} \theta_s^2 ds \right] < \infty. \end{aligned}$$

Remark 3.2.1. We can also compute the relative entropy process $h_t(\mathbb{Q}|\mathbb{P}) := h(\mathbb{Q}_t|\mathbb{P}_t) = \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{2} \int_0^t \theta_s^2 ds \right]$. Observe that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[h_{\tau_d}(\mathbb{Q}|\mathbb{P})] &= \mathbb{E}^{\mathbb{Q}} \left[\int_0^{\infty} \lambda e^{-\lambda t} h_t(\mathbb{Q}|\mathbb{P}) dt \right] = \mathbb{E}^{\mathbb{Q}} \left[\int_0^{\infty} \lambda e^{-\lambda t} \frac{1}{2} \int_0^t \theta_s^2 ds dt \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{2} \int_0^{\infty} \theta_s^2 \int_s^{\infty} \lambda e^{-\lambda t} dt ds \right] = \mathbb{E}^{\mathbb{Q}} \left[\frac{1}{2} \int_0^{\infty} e^{-\lambda s} \theta_s^2 ds \right] = h^d(\mathbb{Q}|\mathbb{P}). \end{aligned}$$

So we can also think of h^d as penalizing the expected relative entropy at death time.

Substituting the expression for $h^d(\mathbb{Q}|\mathbb{P})$ into (3.1), we rewrite the robust value function as:

Definition 3.2.2 (Robust value function).

$$\psi(w; \varepsilon) = \inf_{\pi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_w^{\mathbb{Q}} \left[1_{\{\tau_b < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_d} \theta_s^2 ds \right],$$

where W has \mathbb{Q} -dynamics (3.3).

When $\lambda > 0$, we may use the distribution of τ_d to further write

$$\psi(w; \varepsilon) = \inf_{\pi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_w^{\mathbb{Q}} \left[\int_0^{\infty} e^{-\lambda s} \left(\lambda 1_{\{\tau_b < s\}} - \frac{1}{2\varepsilon} \theta_s^2 \right) ds \right].$$

Denote by ψ_0 the non-robust value function and by \mathbf{p} the robust value function when $\lambda = 0$, i.e. when the individual never dies. ψ_0 has the explicit formula (see [82]):

$$(3.4) \quad \psi_0(w) = \begin{cases} 1, & w \leq b; \\ \left(\frac{c-rw}{c-rb}\right)^d, & b \leq w \leq c/r; \\ 0, & w \geq c/r; \end{cases}$$

and the optimal investment strategy in feedback form is given by

$$\pi_0(w) = \frac{\mu - r}{\sigma^2} \frac{c - rw}{(d - 1)r}$$

for $w \in (b, w_s)$, where

$$(3.5) \quad d = \frac{1}{2r} \left[(r + \lambda + R) + \sqrt{(r + \lambda + R)^2 - 4r\lambda} \right] > 1, \quad R = \frac{1}{2} \left(\frac{\mu - r}{\sigma} \right)^2.$$

Throughout this paper, d and R will be reserved for the constants defined above. We will also provide an explicit formula for \mathbf{p} later. For now, we make the simple observation:

$$(3.6) \quad 0 \leq \psi_0 \leq \psi \leq \mathbf{p} \leq 1,$$

where the second inequality holds because $\mathbb{P} \in \mathcal{Q}$ so that

$$\psi_0 = \inf_{\pi \in \mathcal{A}} \mathbb{P}_w(\tau_b < \tau_d) \leq \inf_{\pi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \left\{ \mathbb{Q}_w(\tau_b < \tau_d) - \frac{1}{\varepsilon} h^d(\mathbb{Q}|\mathbb{P}) \right\} = \psi,$$

the third inequality holds because ruin before death is no more likely than ruin before infinity, and the last inequality holds because we are optimizing a real probability minus a nonnegative penalty. This means we can treat the robust optimal value as a conservative ruin probability. The penalty term will only cause a small distortion on the ruin probability and will never drive it negative because only measures with

small relative entropy are relevant, i.e. have the possibility of being worse than the reference measure.

The definition of $\psi(w; \varepsilon)$ implies it is non-decreasing in ε , since the penalty gets smaller as ε gets larger. We will suppress the argument ε throughout the rest of this paper unless we need to emphasize the ε -dependence. The limit as $\varepsilon \downarrow 0$ gives us the non-robust value function ψ_0 . The limit as $\varepsilon \rightarrow \infty$ gives us the worst-case value function:

$$\psi_\infty(w) := \inf_{\pi \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{Q}_w(\tau_b < \tau_d).$$

For the worse-case problem, the optimal investment strategy is not to invest at all since the drift can be arbitrarily unfavorable (negative if one longs and positive if one shorts) without incurring any penalty. The individual can only hope to “win” the game by dying quickly enough before consumption drags her wealth down to the ruin level. In this case, the agent’s wealth solves the deterministic differential equation:

$$dW_t = (rW_t - c)dt, \quad W_0 = w.$$

Simple computation leads to $\tau_b = \frac{1}{r} \ln \frac{c-rb}{c-rw}$ and $\mathbb{Q}(\tau_b < \tau_d) = e^{-\lambda\tau_b} = \left(\frac{c-rw}{c-rb}\right)^{\frac{\lambda}{r}}$ for $w \in [b, w_s]$ and for all $\mathbb{Q} \in \mathcal{Q}$. So

$$(3.7) \quad \psi_\infty(w) = \left(\frac{c-rw}{c-rb}\right)^{\frac{\lambda}{r}}, \quad w \in [b, w_s].$$

Alternatively, we can obtain the above formula for ψ_∞ by solving (3.12) with ε set to infinity; a verification theorem has to be done then.

Back to the general case. $\psi(w)$ is non-increasing in w since the individual is clearly better off with a larger initial wealth. When $w \leq b$, $\tau_b = 0$ and $\psi(w) = 1$ because the inner supremum can always be attained by the reference measure \mathbb{P} . Notice that by (3.6), we have continuity of ψ at $w = b$ since $1 \geq \lim_{w \rightarrow b} \psi(w) \geq \lim_{w \rightarrow b} \psi_0(w) = 1$.

Let $w_s := c/r$. w_s gives a “safe” wealth level at which the individual can sustain her consumption by putting all her money in the bank and consuming the interest. This means $\psi(w) = 0$ when $w \geq w_s$. Drift uncertainty is irrelevant here since the individual can always play safe by not investing in the risky asset. We also have continuity of ψ at $w = w_s$ because $0 \leq \lim_{w \rightarrow w_s} \psi(w) \leq \lim_{w \rightarrow w_s} \psi_\infty(w) = 0$.

The associated HJB equation for ψ in the interval (b, w_s) is

$$(3.8) \quad \lambda\psi(w) = \inf_{\pi} \sup_{\theta} \left\{ -\frac{1}{2\varepsilon} \theta^2 + (rw - c + (\mu + \sigma\theta - r)\pi) \psi'(w) + \frac{1}{2} \sigma^2 \pi^2 \psi''(w) \right\},$$

with boundary conditions $\psi(b) = 1$ and $\psi(w_s) = 0$. Notice that the expression inside the braces is quadratic in θ with negative leading coefficient. By the first order condition, the optimal θ given π equals $\sigma\varepsilon\pi\psi'$. Substituting $\theta = \sigma\varepsilon\pi\psi'$ back into (3.8), we get

$$(3.9) \quad \lambda\psi = \inf_{\pi} \left\{ \frac{1}{2} \sigma^2 (\varepsilon(\psi')^2 + \psi'') \pi^2 + (\mu - r)\psi'\pi + (rw - c)\psi' \right\}.$$

Suppose $\varepsilon(\psi')^2 + \psi'' > 0$, we use first order condition again to find the candidate optimizer

$$(3.10) \quad \pi^* = -\frac{\mu - r}{\sigma^2} \frac{\psi'}{\varepsilon(\psi')^2 + \psi''}.$$

It follows that

$$(3.11) \quad \theta^* = -\frac{\mu - r}{\sigma} \frac{\varepsilon(\psi')^2}{\varepsilon(\psi')^2 + \psi''}.$$

Substituting (3.10) into (4.4), we obtain the following Dirichlet boundary value problem:

$$(3.12a) \quad \lambda\psi = -\frac{R(\psi')^2}{\varepsilon(\psi')^2 + \psi''} + (rw - c)\psi'$$

$$(3.12b) \quad \psi(b) = 1, \quad \psi(w_s) = 0$$

where R is the positive constant defined in (3.5). When $\varepsilon = 0$, we recover the non-robust value function ψ_0 whose formula is given in (3.4). When $\varepsilon = \infty$, we get the worst-case value function ψ_∞ whose formula is given in (3.7).

Remark 3.2.3. The Isaacs condition does not hold for our robust problem without further restrictions on model parameters. Suppose $\psi'' < 0$ but $\varepsilon(\psi')^2 + \psi'' > 0$, then maximizing over θ first and minimizing over π second in (3.8) will lead to a finite Hamiltonian, but minimizing over π first and maximizing over θ second will lead to an unbounded Hamiltonian. From another perspective, we expect the value function of each fixed-measure lifetime ruin problem to be convex, otherwise the Hamiltonian would explode. Maximizing over these convex functions will yield a convex function. On the other hand, our robust value function may be concave in certain region. When $r > \lambda$, the worst-case value function ψ_∞ is concave. Since $\psi(w; \varepsilon)$ increases to $\psi_\infty(w)$ as $\varepsilon \rightarrow \infty$, $\psi(w; \varepsilon)$ cannot be convex everywhere for ε sufficiently large. See Proposition 3.7.1 for a more detailed discussion on how convexity depends on λ , r and ε .

Rigorous analysis of equation (4.4) will be done in Sections 3.4 and 3.5. Section 3.3 provides an explicit solution to the Dirichlet problem (3.12) when $\lambda = 0$. We end this section with our main result the proof of which is given at the end of Section 3.6.

Theorem 3.2.4. *The robust value function ψ satisfies $\psi(w) = 1$ for $w \leq b$, $\psi(w) = 0$ for $w \geq w_s$. For $w \in (b, w_s)$, $\psi(w)$ is the unique $C^1[b, w_s] \cap C^2(b, w_s)$ solution to (3.8) or (4.4) satisfying the boundary conditions $\psi(b) = 1$ and $\psi(w_s) = 0$. The optimal investment policy is*

$$\pi_t^* = -\frac{\mu - r}{\sigma^2} \frac{\psi'(W_t)}{\varepsilon(\psi'(W_t))^2 + \psi''(W_t)} 1_{(b, w_s)}(W_t),$$

and the optimal drift distortion is $\sigma\theta^*$ where

$$\theta_t^* = -\frac{\mu - r}{\sigma} \frac{\varepsilon(\psi'(W_t))^2}{\varepsilon(\psi'(W_t))^2 + \psi''(W_t)} 1_{(b, w_s)}(W_t).$$

3.3 Explicit solution in the zero-hazard rate case

Setting $\lambda = 0$ in (3.12), we get

$$(3.13) \quad \begin{aligned} 0 &= -\frac{R(\psi')^2}{\varepsilon(\psi')^2 + \psi''} + (rw - c)\psi' \\ \psi(b) &= 1, \quad \psi(w_s) = 0. \end{aligned}$$

Using the exponential transformation $\phi = e^{\varepsilon\psi}$, also called Cole-Hopf transformation in PDE theory, the nonlinearity in the denominator is removed and (3.13) becomes

$$\begin{aligned} 0 &= -R\frac{(\phi')^2}{\phi''} + (rw - c)\phi' \\ \phi(b) &= e^\varepsilon, \quad \phi(w_s) = 1. \end{aligned}$$

Suppose $\phi' \neq 0$ and let $u = \phi'$. The second order ordinary differential equation (ODE) is further reduced to

$$u' = \frac{R}{rw - c}u,$$

the general solution of which is given by

$$u(w) = Ae^{R\int_b^w \frac{1}{rz-c}dz} = A\left(\frac{c-rw}{c-rb}\right)^{\frac{R}{r}}, \quad A \in \mathbb{R}.$$

It follows that

$$\phi(w) = e^\varepsilon + A\int_b^w \left(\frac{c-rz}{c-rb}\right)^{\frac{R}{r}} dz = e^\varepsilon - A\frac{c-rb}{R+r} \left[\left(\frac{c-rw}{c-rb}\right)^{\frac{R}{r}+1} - 1 \right].$$

Using the boundary condition at the safe level, we can determine the constant A and obtain

$$\phi(w) = 1 + (e^\varepsilon - 1) \left(\frac{c-rw}{c-rb}\right)^{\frac{R}{r}+1}.$$

So the solution to the Dirichlet problem (3.13) is

$$(3.14) \quad \psi(w) = \frac{1}{\varepsilon} \ln \left[1 + (e^\varepsilon - 1) \left(\frac{c - rw}{c - rb} \right)^{\frac{R}{r} + 1} \right].$$

The feedback forms (3.10), (3.11) become

$$\varpi = \frac{2(c - rw)}{\mu - r},$$

$$\vartheta = -\frac{2\sigma(R + r)}{\mu - r} \frac{(e^\varepsilon - 1) \left(\frac{c - rw}{c - rb} \right)^{\frac{R}{r} + 1}}{1 + (e^\varepsilon - 1) \left(\frac{c - rw}{c - rb} \right)^{\frac{R}{r} + 1}}.$$

The solution given by (3.14) is a $C^1[b, w_s] \cap C^2[b, w_s]$ function. ϖ and ϑ are bounded, Lipschitz continuous functions of the state variable on $[b, w_s]$. So a verification theorem can be easily done, showing the function given by (3.14) is indeed the robust value function \mathbf{p} on the interval $[b, w_s]$, and ϖ, ϑ are the optimal feedback controls. We summarize the results in the following theorem.

Theorem 3.3.1. *When $\lambda = 0$, the robust value function is given by*

$$\mathbf{p}(w) = \frac{1}{\varepsilon} \ln \left[1 + (e^\varepsilon - 1) \left(\frac{c - rw}{c - rb} \right)^{\frac{R}{r} + 1} \right]$$

for $b \leq w \leq w_s$, $\mathbf{p}(w) = 0$ for $w \leq b$ and $\mathbf{p}(w) = 1$ for $w \geq w_s$. The optimal investment policy is

$$\varpi_t = \frac{2(c - rW_t)}{\mu - r} 1_{(b, w_s)}(W_t),$$

and the optimal drift distortion is $\sigma\vartheta$ where

$$\vartheta_t = -\frac{2\sigma(R + r)}{\mu - r} \frac{(e^\varepsilon - 1) \left(\frac{c - rW_t}{c - rb} \right)^{\frac{R}{r} + 1}}{1 + (e^\varepsilon - 1) \left(\frac{c - rW_t}{c - rb} \right)^{\frac{R}{r} + 1}} 1_{(b, w_s)}(W_t).$$

One observation is the loss of convexity of the value function compared with the non-robust case. This is caused by the nonlinear term $\varepsilon(\psi')^2$. When ε is zero, ψ'' must be non-negative (in fact, strictly positive if $\psi' \neq 0$) for the Hamiltonian in (4.4) to be

finite. When ε is nonzero, ψ'' is allowed to take negative values as long as $\varepsilon(\psi')^2 + \psi''$ is non-negative. The larger the ε , the more concave the value function could potentially be. Another interesting feature is that when hazard rate is zero, the pre-ruin optimal investment policy is independent of both the ambiguity aversion parameter ε and the ruin level b . Also, we see that for $w \in (b, w_s)$, $\lim_{\varepsilon \rightarrow \infty} \vartheta(w) = -\frac{2\sigma(R+r)}{\mu-r}$. In terms of the optimally distorted Sharpe ratio, we have

$$\lim_{\varepsilon \rightarrow \infty} \left(\frac{\mu - r}{\sigma} + \vartheta(w) \right) = -\frac{2\sigma r}{\mu - r}.$$

Figure 3.1 shows plots for the robust ruin probability \mathbf{p} and the optimally distorted Sharpe ratio $\frac{\mu-r}{\sigma} + \vartheta$ with parameters $c = 1$, $b = 1$, $r = 0.02$, $\mu = 0.1$, $\sigma = 0.15$ and $\varepsilon = 0, 1, 5, 10, 50$. We leave out the plot for ϖ since it is a simple downward sloping linear function, and is independent of ε . It is worth mentioning that $\varpi \geq \pi_0$ (in fact, ϖ dominates the optimal robust policy with death by Lemma 3.6.1), i.e. the individual adopts a more aggressive investment strategy when life is perpetual.

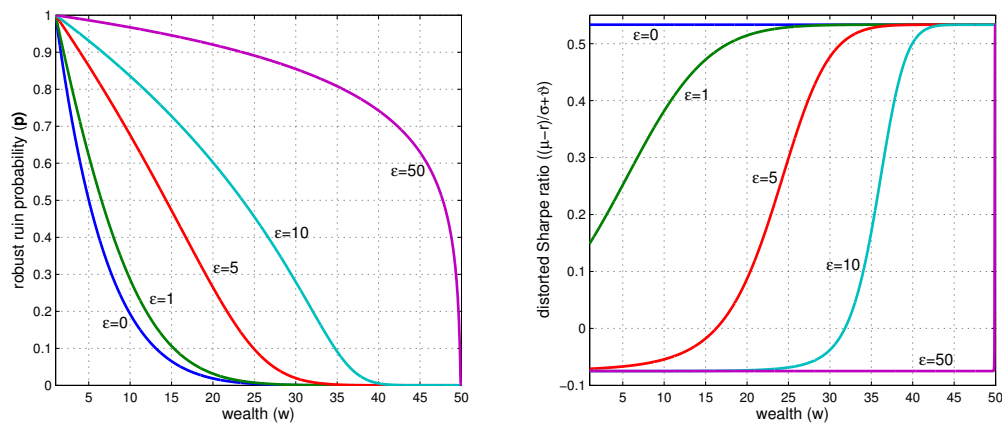


Figure 3.1: Robust ruin probabilities and distorted Sharpe ratios when $\lambda = 0$.

Before we move on to the general case, let us make one more remark regarding differentiability at the safe level.

Remark 3.3.2. From the explicit formula for \mathfrak{p} , we see that \mathfrak{p} has zero derivative at the safe level. Since \mathfrak{p} bounds any general ψ from above, this property is also shared by ψ . Indeed,

$$0 \geq \lim_{w \rightarrow w_s^-} \frac{\psi(w)}{w - w_s} \geq \lim_{w \rightarrow w_s^-} \frac{\mathfrak{p}(w)}{w - w_s} = \mathfrak{p}'(w_s) = 0.$$

3.4 Existence and uniqueness of viscosity solution

Our goal in this section is to show the nonlinear degenerate elliptic Dirichlet problem

$$(3.15a) \quad F(w, u, u', u'') = 0,$$

$$(3.15b) \quad u(b) = 1, \quad u(w_s) = 0,$$

where

$$F(w, u, u', u'') := \lambda u - \inf_{\pi} \left\{ \frac{1}{2} \sigma^2 (\varepsilon(u')^2 + u'') \pi^2 + (\mu - r) u' \pi + (rw - c) u' \right\}$$

has a unique viscosity solution satisfying certain properties. Notice that F can be written as the supremum of a family of continuous functions, hence is lower semi-continuous (l.s.c.). The strategy is to first prove a comparison principle, and then use Perron's Method introduced by Ishii [46] (also described in [29]) to construct a viscosity solution as the supremum over a class of viscosity subsolutions.

3.4.1 Comparison principle

The proof of the comparison principle uses a standard doubling of variable technique together with Crandall-Ishii's lemma. The classical argument is slightly modified to take care of the unboundedness of the control space. It turns out, luckily, that the nonlinear term $\varepsilon(u')^2$ does not add any difficulty.

Proposition 3.4.1. *Let u, v be an upper semi-continuous (u.s.c.) viscosity subsolution and a l.s.c. viscosity supersolution of $F = 0$, respectively. Suppose u, v are bounded, and either $v \geq 0$ or $u > 0$ in (b, w_s) . If $u \leq v$ on $\partial(b, w_s)$, then $u \leq v$ on $[b, w_s]$.*

Proof. Suppose that, on the contrary, $\delta := \sup_{x \in (b, w_s)} (u - v)(x) > 0$. $\delta < \infty$ since u and v are assumed to be bounded. By the upper semi-continuity of $u - v$, there exists $x^* \in (b, w_s)$ such that $u(x^*) - v(x^*) = \delta$. For every $\alpha > 0$, define

$$\Psi_\alpha(x, y) := u(x) - v(y) - \frac{\alpha}{2}|x - y|^2.$$

It is clear that $\sup_{x, y \in (b, w_s)} \Psi_\alpha(x, y) \geq \delta$ since we can always choose $x = y$. By the upper semi-continuity of $u(x) - v(y)$, there exists $\hat{x}_\alpha, \hat{y}_\alpha$ such that $\sup_{x, y \in (b, w_s)} \Psi_\alpha(x, y) = \Psi_\alpha(\hat{x}_\alpha, \hat{y}_\alpha)$. We have

$$u(x^*) - v(x^*) \leq u(\hat{x}_\alpha) - v(\hat{y}_\alpha) - \frac{\alpha}{2}|\hat{x}_\alpha - \hat{y}_\alpha|^2.$$

This implies

$$(3.16) \quad \frac{\alpha}{2}|\hat{x}_\alpha - \hat{y}_\alpha|^2 \leq u(\hat{x}_\alpha) - v(\hat{y}_\alpha) - (u(x^*) - v(x^*)).$$

Since $[b, w_s]$ is compact, we can find a sequence $\alpha_n \rightarrow \infty$ such that $(\hat{x}_n, \hat{y}_n) := (\hat{x}_{\alpha_n}, \hat{y}_{\alpha_n})$ converges to (\hat{x}, \hat{y}) as $n \rightarrow \infty$. Replacing α by α_n and letting $n \rightarrow \infty$ in (4.35), we obtain

$$(3.17) \quad \begin{aligned} \limsup_n \frac{\alpha_n}{2}|\hat{x}_n - \hat{y}_n|^2 &\leq \limsup_n (u(\hat{x}_n) - v(\hat{y}_n)) - (u(x^*) - v(x^*)) \\ &\leq u(\hat{x}) - v(\hat{y}) - (u(x^*) - v(x^*)), \end{aligned}$$

where the second inequality is due to the upper semi-continuity of $u(x) - v(y)$. Since the right hand side of (4.36) is finite and $\alpha_n \rightarrow \infty$, we must have $\hat{x} = \hat{y}$, and (4.36) yields

$$0 \leq \limsup_n \frac{\alpha_n}{2}|\hat{x}_n - \hat{y}_n|^2 \leq u(\hat{x}) - v(\hat{x}) - (u(x^*) - v(x^*)) \leq 0,$$

which implies $u(\hat{x}) - v(\hat{x}) = u(x^*) - v(x^*) = \delta$, $\alpha_n |\hat{x}_n - \hat{y}_n|^2 \rightarrow 0$ and

$$(3.18) \quad \delta \leq \sup_{x, y \in (b, w_s)} \Psi_{\alpha_n}(x, y) = u(\hat{x}_n) - v(\hat{y}_n) - \frac{\alpha_n}{2} |\hat{x}_n - \hat{y}_n|^2 \rightarrow u(x^*) - v(x^*) = \delta$$

as $n \rightarrow \infty$. Now, since $u \leq v$ on $\partial(b, w_s)$, we must have $\hat{x} \in (b, w_s)$. So $\hat{x}_n, \hat{y}_n \in (b, w_s)$ for sufficiently large n . By Crandall-Ishii's lemma, we can find sequences A_n, B_n satisfying $-3\alpha_n \leq A_n \leq B_n \leq 3\alpha_n$ and

$$(\alpha_n(\hat{x}_n - \hat{y}_n), A_n) \in \bar{J}_{(b, w_s)}^{2,+} u(\hat{x}_n), \quad (\alpha_n(\hat{x}_n - \hat{y}_n), B_n) \in \bar{J}_{(b, w_s)}^{2,-} v(\hat{y}_n),$$

where $\bar{J}_{(b, w_s)}^{2,+} u(\hat{x}_n), \bar{J}_{(b, w_s)}^{2,-} v(\hat{y}_n)$ are the closure of the second order superjet and subjet, respectively. Since u is a viscosity subsolution of $F = 0$ and F is l.s.c., we have by [80, Proposition 6.11.i] that

$$(3.19) \quad F(\hat{x}_n, u(\hat{x}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n) \leq 0.$$

The finiteness of $F(\hat{x}_n, u(\hat{x}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n)$ implies either $\varepsilon \alpha_n^2 (\hat{x}_n - \hat{y}_n)^2 + A_n > 0$ or $\varepsilon \alpha_n^2 (\hat{x}_n - \hat{y}_n)^2 + A_n = \alpha_n(\hat{x}_n - \hat{y}_n) = 0$. We consider each case separately.

Case 1. $\varepsilon \alpha_n^2 (\hat{x}_n - \hat{y}_n)^2 + A_n > 0$. In this case, we also have $\varepsilon \alpha_n^2 (\hat{x}_n - \hat{y}_n)^2 + B_n > 0$. Since $F(w, u, u', u'')$ is continuous in the region $\varepsilon(u')^2 + u'' > 0$, the supersolution property of v implies

$$F(\hat{y}_n, v(\hat{y}_n), \alpha_n(\hat{x}_n - \hat{y}_n), B_n) \geq 0.$$

(See [80, Proposition 6.11.ii].) So we have

$$(3.20) \quad F(\hat{x}_n, u(\hat{x}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n) \leq 0 \leq F(\hat{y}_n, v(\hat{y}_n), \alpha_n(\hat{x}_n - \hat{y}_n), B_n) < \infty.$$

Using the expression of F , we obtain from (4.37) and (3.20) that

$$\begin{aligned}
\lambda\delta &\leq \lambda(u(\hat{x}_n) - v(\hat{y}_n)) \\
&= F(\hat{x}_n, u(\hat{x}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n) - F(\hat{x}_n, v(\hat{y}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n) \\
&\leq F(\hat{y}_n, v(\hat{y}_n), \alpha_n(\hat{x}_n - \hat{y}_n), B_n) - F(\hat{x}_n, v(\hat{y}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n) \\
&= \frac{R\alpha_n^2(\hat{x}_n - \hat{y}_n)^2(A_n - B_n)}{[\varepsilon\alpha_n^2(\hat{x}_n - \hat{y}_n)^2 + B_n][\varepsilon\alpha_n^2(\hat{x}_n - \hat{y}_n)^2 + A_n]} + r\alpha_n(\hat{x}_n - \hat{y}_n)^2 \\
&\leq r\alpha_n(\hat{x}_n - \hat{y}_n)^2.
\end{aligned}$$

Letting $n \rightarrow \infty$, we arrive at the contradiction $\lambda\delta \leq 0$.

Case 2. $\varepsilon\alpha_n^2(\hat{x}_n - \hat{y}_n)^2 + A_n = \alpha_n(\hat{x}_n - \hat{y}_n) = 0$. In this case, Equation (4.38) reads

$$\lambda u(\hat{x}_n) = F(\hat{x}_n, u(\hat{x}_n), \alpha_n(\hat{x}_n - \hat{y}_n), A_n) \leq 0.$$

If u is strictly positive, this cannot happen. So assume we are in the case where v is non-negative. But this implies $u(\hat{x}_n) - v(\hat{x}_n) \leq 0$, contradicting $u(\hat{x}_n) - v(\hat{y}_n) \geq \sup_{x,y \in (b,w_s)} \Psi_{\alpha_n}(x,y) \geq \delta$. \square

Corollary 3.4.2. *There is at most one viscosity solution to the Dirichlet problem (3.15) that is bounded, non-negative, and continuous at the boundary.*

3.4.2 Perron's method

We mimic the proof of [29, Theorem 4.1], and begin with a max-stability result on the set of viscosity subsolutions.

Lemma 3.4.3. *Let \mathcal{U} be a non-empty family of u.s.c. viscosity subsolutions of $F = 0$.*

Define

$$\mathbf{u}(w) = \sup_{u \in \mathcal{U}} u(w).$$

Let \mathbf{u}^ be the u.s.c. envelope of \mathbf{u} and assume $\mathbf{u}^*(w) < \infty$ for $w \in (b, w_s)$. Then \mathbf{u}^* is a viscosity subsolution of $F = 0$.*

Proof. By Lemma 4.2 of [29]. Note that although their function F is \mathbb{R} -valued, the proof works exactly the same way when F is allowed to take ∞ as a value as long as it is l.s.c. which is satisfied in our case. \square

Next, we need to find an (u.s.c.) viscosity subsolution whose l.s.c. envelope satisfies the boundary conditions (3.15b), and a (l.s.c.) viscosity supersolution whose u.s.c. envelope satisfies the boundary conditions (3.15b). Obviously, we should aim at those functions that bound the robust value function from below and above, and we have two natural candidates: ψ_0 and \mathbf{p} .⁶ Indeed,

$$\begin{aligned} F(w, \psi_0, \psi'_0, \psi''_0) &= \lambda\psi_0 - \inf_{\pi} \left\{ \frac{1}{2}\sigma^2 (\varepsilon(\psi'_0)^2 + \psi''_0) \pi^2 + (\mu - r)\psi'_0\pi + (rw - c) \psi'_0 \right\} \\ &= \lambda\psi_0 + \frac{R(\psi'_0)^2}{\varepsilon(\psi'_0)^2 + \psi''_0} - (rw - c) \psi'_0 \\ &\leq \lambda\psi_0 + \frac{R(\psi'_0)^2}{\psi''_0} - (rw - c) \psi'_0 = 0, \end{aligned}$$

where in the second equality we used $\psi''_0 > 0$ in (b, w_s) , and

$$\begin{aligned} F(w, \mathbf{p}, \mathbf{p}', \mathbf{p}'') &= \lambda\mathbf{p} - \inf_{\pi} \left\{ \frac{1}{2}\sigma^2 (\varepsilon(\mathbf{p}')^2 + \mathbf{p}'') \pi^2 + (\mu - r)\mathbf{p}'\pi + (rw - c) \mathbf{p}' \right\} \\ &= \lambda\mathbf{p} \geq 0. \end{aligned}$$

Remark 3.4.4. If these natural candidates were not available, we could start with the constant subsolution $u \equiv 0$ (resp. supersolution $v \equiv 1$), and modify it near the ruin level (resp. safe level) by a construction similar to that on page 25 of [29] so that the boundary conditions are satisfied.

Proposition 3.4.5 (Perron's method). *There exists a continuous viscosity solution to the Dirichlet problem (3.15) that takes values in $[0, 1]$. More precisely, it is bounded from below by ψ_0 and from above by \mathbf{p} .*

⁶We can also use ψ_∞ as the upper bound.

Proof. Let $\underline{u} = \psi_0$ and $v = \mathbf{p}$. Both are $[0, 1]$ -valued continuous functions. Define

$$(3.21) \quad \mathbf{u}(w) := \sup\{u(w) : \underline{u} \leq u \leq v \text{ and } u \text{ is an u.s.c. subsolution of } F = 0\}.$$

For any function u , denote by u^* and u_* its u.s.c. envelope and l.s.c. envelope, respectively. We have $\underline{u} = \underline{u}_* \leq \mathbf{u}_* \leq \mathbf{u} \leq \mathbf{u}^* \leq v^* = v$. Since \underline{u} and v agree on the boundary, we know \mathbf{u} is continuous at the boundary and satisfies the boundary condition (3.15b). Since $\mathbf{u}^* \leq v < \infty$, Lemma 3.4.3 implies \mathbf{u}^* is a viscosity subsolution of $F = 0$. If we can show \mathbf{u}_* is a viscosity supersolution of $F = 0$, we can then apply comparison principle to get $\mathbf{u}^* \leq \mathbf{u}_*$, and conclude that \mathbf{u} is a continuous viscosity solution to the Dirichlet problem (3.15). The rest is devoted to the proof of the supersolution property of \mathbf{u}_* .

Suppose \mathbf{u}_* is not a viscosity supersolution of $F = 0$. Then there exists $w_0 \in (b, w_s)$ and $\varphi \in C^2(b, w_s)$ such that $\mathbf{u}_* - \varphi$ has a strict minimum zero at w_0 and $F(w_0, \varphi(w_0), \varphi'(w_0), \varphi''(w_0)) < 0$. Here $F < \infty$ implies either $\varepsilon(\varphi'(w_0))^2 + \varphi''(w_0) > 0$ or $\varphi''(w_0) = \varphi'(w_0) = 0$. In the latter case, we get $\mathbf{u}_*(w_0) = \varphi(w_0) < 0$ which cannot happen because $\mathbf{u}_* \geq \underline{u} \geq 0$. So we are in the former case. By continuity of F in the region $\varepsilon(u')^2 + u'' > 0$, there exists $\delta, \gamma > 0$ such that $F(w, \varphi(w) + \gamma, \varphi'(w), \varphi''(w)) < 0$ for all $w \in B_\delta(w_0) \subset \overline{B}_\delta(w_0) \subset (b, w_s)$. Let $\varphi_\gamma(w) := \varphi(w) + \gamma$. Then φ_γ is a classical subsolution of $F = 0$ in $B_\delta(w_0)$. Since $\mathbf{u}_* > \varphi$ in $(b, w_s) \setminus \{w_0\}$, we can choose γ small so that $\mathbf{u}_* > \varphi + \gamma = \varphi_\gamma$ on $\partial B_\delta(w_0)$. Define

$$U := \begin{cases} \mathbf{u}^* \vee \varphi_\gamma & \text{in } B_\delta(w_0), \\ \mathbf{u}^* & \text{otherwise.} \end{cases}$$

Since $\mathbf{u}^* < \infty$ and $\varphi_\gamma \leq \mathbf{u}_* + \gamma < \infty$ in $B_\delta(w_0)$, by Lemma 3.4.3, U^* is a viscosity subsolution of $F = 0$. Since $U^* = \mathbf{u}^* \leq v$ on $\partial(b, w_s)$, comparison principle (Proposition 3.4.1) implies $U^* \leq v$ on $[b, w_s]$. So U^* belongs to the set on the right hand

side of (3.21), and thus $\mathbf{u}^* \leq U \leq U^* \leq \mathbf{u} \leq \mathbf{u}^*$, where the second last inequality is due to the maximality of \mathbf{u} . Therefore, we obtain $U = \mathbf{u}^*$.

On the other hand, by the definition of the semi-continuous envelope, there exists a sequence $(w_n) \subset B_\delta(w_0)$ such that $w_n \rightarrow w_0$ and $\mathbf{u}^*(w_n) \rightarrow \mathbf{u}_*(w_0)$. It follows that $\varphi_\gamma(w_n) - \mathbf{u}^*(w_n) = \varphi(w_n) + \gamma - \mathbf{u}^*(w_n) \rightarrow \gamma > 0$. So for n sufficiently large, $U(w_n) = \varphi_\gamma(w_n) > \mathbf{u}^*(w_n) + \gamma/2$. We get a contradiction. This completes the proof that \mathbf{u}_* is a viscosity supersolution of $F = 0$. \square

Up to this point, we have established the existence and uniqueness of a continuous viscosity solution to the Dirichlet problem (3.15). Denote this solution by \hat{u} . We have $\psi_0 \leq \hat{u} \leq \mathbf{p}$. The next goal is to upgrade regularity.

3.5 Regularity

3.5.1 An equivalent convex problem

One difficulty of directly proving regularity for problem (3.15) is the lack of convexity of \hat{u} caused by the nonlinear term $\varepsilon(u')^2$. Motivated by how we solved the $\lambda = 0$ case, we use the Cole-Hopf transformation $v = e^{\varepsilon u}$ to obtain an equivalent convex problem:

$$(3.22a) \quad G(w, v, v', v'') = 0,$$

$$(3.22b) \quad v(b) = e^\varepsilon, \quad v(w_s) = 1.$$

where

$$G(w, v, v', v'') := \lambda v \ln v - \inf_{\pi} \left\{ \frac{1}{2} \sigma^2 v'' \pi^2 + (\mu - r) v' \pi + (rw - c) v' \right\}.$$

The solution to the transformed problem is expected to be convex, otherwise G would explode. Although (3.22a) is only understood in viscosity sense for now, one can

expect, intuitively, if at every interior point, every test function above the viscosity solution (for the subsolution property) is convex in a neighborhood of that point, then the viscosity solution should be convex as well.

Since we already have a continuous viscosity solution \hat{u} of problem (3.15), it can be easily verified that $\hat{v} := e^{\varepsilon\hat{u}}$ is a continuous viscosity solution of problem (3.22) satisfying $e^{\varepsilon\psi_0} \leq \hat{v} \leq e^{\varepsilon p}$. Moreover, the comparison principle for (3.15) immediately yields a comparison principle for (3.22). We summarize these results in the following two lemmas.

Lemma 3.5.1. *Let u, v be strictly positive u.s.c. viscosity subsolution and l.s.c. viscosity supersolution of (3.22a), respectively. Suppose u, v are bounded and bounded away from zero and either $u > 1$ or $v \geq 1$ in (b, w_s) . If $u \leq v$ on $\partial(b, w_s)$, then $u \leq v$ on $[b, w_s]$.*

Proof. It is easy to check $\frac{1}{\varepsilon} \ln u$ (resp. $\frac{1}{\varepsilon} \ln v$) is an u.s.c. subsolution (resp. a l.s.c. supersolution) of (3.15) satisfying all assumptions of Proposition 3.4.1. \square

Lemma 3.5.2. *$\hat{v} := e^{\varepsilon\hat{u}}$ is the unique (continuous) viscosity solution to the Dirichlet problem (3.22) among all viscosity solutions that are bounded, continuous at the boundary and satisfy $v \geq 1$ in (b, w_s) . Moreover, $e^{\varepsilon\psi_0} \leq \hat{v} \leq e^{\varepsilon p}$.*

3.5.2 From convexity to C^∞

In the subsection, we prove regularity for the solution \hat{v} of the transformed problem. The regularity for $\hat{u} = \frac{1}{\varepsilon} \ln \hat{v}$ immediately follows.

Lemma 3.5.3. *\hat{v} is strictly convex and strictly decreasing on $[b, w_s]$.*

Proof. First, let us show (non-strict) interior convexity. Suppose \hat{v} is not convex in (b, w_s) . Then by [4, Lemma 1], there exists $w_0 \in (b, w_s)$ and $(p, A) \in J^{2,+}\hat{v}(w_0)$ with

$A < 0$. We therefore have $G(w_0, \hat{v}(w_0), p, A) = \infty$. But by the semi-jets formulation of viscosity solution (see e.g. [80, Proposition 6.11.i]), we have $G(w_0, \hat{v}(w_0), p, A) \leq 0$. We get a contradiction. Since \hat{v} is continuous, interior convexity can be extended to the boundary.

The convexity of \hat{v} implies its left and right derivatives $D^\pm \hat{v}$ exists (in \mathbb{R} for interior points and in $\mathbb{R} \cup \{\pm\infty\}$ for boundary points) and are non-decreasing.⁷ Since we have showed $0 \leq \hat{u} \leq \mathfrak{p}$ and we know $\mathfrak{p}'_0(w_s) = 0$, an argument exactly the same as Remark 3.3.2 yields $D^- \hat{u}(w_s) = 0$. It follows that $D^- \hat{v}(w_s) = \varepsilon D^- \hat{u}(w_s) \hat{v}(w_s) = 0$. So $D^\pm \hat{v} \leq \hat{v}'(w_s) = 0$. Suppose $D^+ \hat{v}(w_0) = 0$ for some $w_0 \in [b, w_s)$ (same if $D^- \hat{v}(w_0) = 0$). Then by monotonicity of $D^+ \hat{v}$, $D^+ \hat{v}(w) = 0 \ \forall w \in [w_0, w_s)$. By convexity,

$$0 = D^+ \hat{v}(w) \leq \frac{\hat{v}(w) - \hat{v}(w_s)}{w - w_s} = \frac{\hat{v}(w) - 1}{w - w_s} \leq 0 \quad \forall w \in [w_0, w_s).$$

We deduce $\hat{v} \equiv 1$ on $[w_0, w_s]$, contradicting the property that $\hat{v} \geq e^{\varepsilon\psi_0} > 1$ in (b, w_s) (see Lemma 3.5.2). Therefore, we must have $D^\pm \hat{v} < 0$ in $[b, w_s)$ which implies \hat{v} is strictly decreasing.

Finally, if \hat{v} is convex but not strictly convex, then it is linear in some open interval $(x, y) \subset (b, w_s)$. Since \hat{v} is strictly decreasing, the line has non-zero slope, say p . But this cannot happen because $G(w, \hat{v}(w), p, 0)$ is unbounded. \square

Being a convex function, \hat{v} has many nice regularity properties. It is differentiable almost everywhere (a.e.), and even twice differentiable a.e. by Alexandroff's classical result [2]. To show C^2 -regularity, we first show C^1 -regularity using properties of viscosity solution and then upgrade to C^2 by analyzing a Poisson equation with the non-homogeneous term expressed in terms of \hat{v} and its first derivative.

⁷Here and in the sequel, at the left (resp. right) boundary point, $D^\pm \hat{v}$ only refers to the right (resp. left) derivative.

Lemma 3.5.4. $(rw - c)D^\pm \hat{v} - \lambda \hat{v} \ln \hat{v}$ is non-negative for all $w \in (b, w_s)$, and strictly positive if w is a point of twice differentiability of \hat{v} .

Proof. By Lemma 2 in [4], $G(w, \hat{v}(w), \hat{v}'(w), \hat{v}''(w)) \leq 0$ at every point $w \in (b, w_s)$ of twice differentiability. Here we note that their lemma is stated for continuous G , but it can be easily modify to accommodate our l.s.c. G . Let $w \in (b, w_s)$ be a point where \hat{v} is twice differentiable. Since $\hat{v}'(w) < 0$, $G(w, \hat{v}(w), \hat{v}'(w), \hat{v}''(w)) \leq 0$ implies $\hat{v}''(w) > 0$, and

$$\lambda \hat{v}(w) \ln \hat{v}(w) + R \frac{(\hat{v}'(w))^2}{\hat{v}''(w)} - (rw - c)\hat{v}'(w) \leq 0.$$

We get

$$(rw - c)\hat{v}'(w) - \lambda \hat{v}(w) \ln \hat{v}(w) \geq R \frac{(\hat{v}'(w))^2}{\hat{v}''(w)} > 0.$$

For arbitrary $w \in (b, w_s)$, since \hat{v} is twice differentiable a.e., we can find a sequence of twice differentiability points $(w_n) \subset (b, w_s)$ which converges to w from the right.

Using the monotonicity of $D^\pm \hat{v}$, we have

$$(rw_n - c)D^\pm \hat{v}(w) \geq (rw_n - c)\hat{v}'(w_n) > \lambda \hat{v}(w_n) \ln \hat{v}(w_n)$$

We are done by letting $n \rightarrow \infty$ and using the continuity of \hat{v} . □

Lemma 3.5.5. $\hat{v} \in C^1[b, w_s]$.

Proof. We first show interior C^1 -regularity. It suffices to show \hat{v} is differentiable since a convex differentiable function is continuously differentiable. Suppose on the contrary, $D^- \hat{v}(w_0) \neq D^+ \hat{v}(w_0)$ at some point $w_0 \in (b, w_s)$. Let $p \in (D^- \hat{v}(w_0), D^+ \hat{v}(w_0))$ and $\epsilon > 0$. The function

$$\varphi(w) = \hat{v}(w_0) + p(w - w_0) + \frac{1}{2\epsilon}(w - w_0)^2$$

satisfies $\hat{v} - \varphi$ has a local minimum at w_0 . By supersolution property of \hat{v} , we get

$$G(w_0, \varphi(w_0), \varphi'(w_0), \varphi''(w_0)) = \lambda \hat{v}(w_0) \ln \hat{v}(w_0) + R\epsilon p^2 - (rw_0 - c)p \geq 0.$$

Since ϵ is arbitrary, we get $\lambda \hat{v}(w_0) \ln \hat{v}(w_0) - (rw_0 - c)p \geq 0$. In view of Lemma 3.5.4, we must have $\lambda \hat{v}(w_0) \ln \hat{v}(w_0) - (rw_0 - c)p = 0$. But this cannot hold for every p . So the subdifferential at every point must be a singleton.

Since \hat{v} is convex on $[b, w_s]$, to extend C^1 -regularity up to the boundary, we only need to check $D^+\hat{v}(b) > -\infty$ and $D^-\hat{v}(w_s) < \infty$. We have already seen in the proof of Lemma 3.5.3 that $D^-\hat{v}(w_s) = 0$. To bound $D^+\hat{v}(b)$ from below, we make use of the derivative of ψ_0 . Simply observe that $D^+\hat{v}(b) = \varepsilon D^+\hat{u}(b)\hat{v}(b)$, and

$$D^+\hat{u}(b) = \lim_{w \rightarrow b^+} \frac{\hat{u}(w) - 1}{w - b} \geq \lim_{w \rightarrow b^+} \frac{\psi_0(w) - 1}{w - b} = D^+\psi_0(b) > -\infty.$$

□

Proposition 3.5.6. $\hat{v} \in C^2(b, w_s)$ and satisfies $\hat{v}' < 0$ and $\hat{v}'' > 0$ in (b, w_s) .⁸ In addition, \hat{v} solves the second order equation

$$(3.23) \quad \lambda v \ln v = -R \frac{(v')^2}{v''} + (rw - c)v', \quad w \in (b, w_s).$$

Proof. $\hat{v}' < 0$ is due to Lemma 3.5.3. Let $f(w) := (rw - c)\hat{v}'(w) - \lambda \hat{v}(w) \ln \hat{v}(w)$. By Lemmas 3.5.4 and 3.5.5, f is continuous, non-negative and a.e. strictly positive in (b, w_s) . Let $g(w) := R(\hat{v}'(w))^2/f(w)$. The proof of C^2 -regularity consists of two steps.

Step 1. Show that for any interval $[w_1, w_2] \subset [b, w_s]$ such that $f > 0$ on $[w_1, w_2]$, $\hat{v} \in C^2[w_1, w_2]$. Notice that g is continuous on $[w_1, w_2]$.

First of all, we show \hat{v} is a viscosity solution of

$$(3.24) \quad -v''(w) + g(w) = 0, \quad w \in (w_1, w_2).$$

⁸The derivatives at $w = b$ is understood to be the continuous extension of interior derivatives.

Let $w_0 \in (w_1, w_2)$ and $\varphi \in C^2(w_1, w_2)$ be any test functions such that $\hat{v} - \varphi$ has a local maximum at w_0 . Since \hat{v} is a C^1 subsolution of $G = 0$, we have $\varphi'(w_0) = \hat{v}'(w_0)$ and

$$G(w_0, \hat{v}(w_0), \hat{v}'(w_0), \varphi''(w_0)) \leq 0.$$

Since $\hat{v}'(w_0) < 0$, we must have $\varphi''(w_0) > 0$ for the above G to be finite. Writing out the expression for G and optimizing over π , we get

$$-f(w_0) + R \frac{(\hat{v}'(w_0))^2}{\varphi''(w_0)} = \lambda \hat{v}(w_0) \ln \hat{v}(w_0) + R \frac{(\hat{v}'(w_0))^2}{\varphi''(w_0)} - (rw_0 - c)\hat{v}'(w_0) \leq 0,$$

which, after multiplying by the positive quantity $\frac{\varphi''(w_0)}{f(w_0)}$, is precisely

$$-\varphi''(w_0) + g(w_0) \leq 0.$$

This shows \hat{v} is a subsolution of (3.24). Let $w_0 \in (w_1, w_2)$ and $\varphi \in C^2(w_1, w_2)$ be any test function such that $\hat{v} - \varphi$ has a local minimum at w_0 . If $\varphi''(w_0) \leq 0$, then we immediately have $-\varphi''(w_0) + g(w_0) \geq 0$ since g is nonnegative. If $\varphi''(w_0) > 0$, then we use \hat{v} is a C^1 supersolution of $G = 0$ to obtain $\varphi'(w_0) = \hat{v}'(w_0)$ and

$$G(w_0, \hat{v}(w_0), \hat{v}'(w_0), \varphi''(w_0)) \geq 0.$$

Optimizing over π in the expression for G , we also get $-\varphi''(w_0) + g(w_0) \geq 0$. This shows \hat{v} is a supersolution of (3.24).

Next, we follow the argument on page 652 of [73] and consider the Poisson equation

$$(3.25) \quad -v'' + g = \epsilon$$

with Dirichlet boundary conditions $v(w_1) = \hat{v}(w_1)$, $v(w_2) = \hat{v}(w_2)$. Here ϵ is a real number of our choice. We can integrate $g - \epsilon$ twice to get a $C^2[w_1, w_2]$ solution, denoted by v_ϵ . To compare \hat{v} with v_ϵ , first take $\epsilon > 0$ and suppose $\hat{v} - v_\epsilon$ has a local

maximum at some point $w_0 \in (w_1, w_2)$. Since \hat{v} is a viscosity subsolution of (3.24), we have

$$-v_\epsilon''(w_0) + g(w_0) \leq 0,$$

which contradicts (3.25). So the maximum must be attained on the boundary where it is zero. This means $\hat{v} \leq v_\epsilon$. Letting $\epsilon \rightarrow 0$ yields $\hat{v} \leq v_0$. The reverse inequality is obtained by taking $\epsilon < 0$ and using \hat{v} is a viscosity supersolution of (3.24). This finishes the proof that $\hat{v} = v_0 \in C^2[w_1, w_2]$.

Step 2. Show $f(w) > 0$ for any $w \in [b, w_s)$.

We use an argument similar to that on page 811-812 of [51]. Pick any point $w_1 \in (b, w_s)$ where $f(w_1) > 0$. Since f is continuous, $f > 0$ in a neighborhood of w_1 . Suppose f vanishes at some point to the left of w_1 . Let $w_0 := \sup\{w \in [b, w_1) : f(w_0) = 0\}$. By step 1, \hat{v} satisfies equation (3.24) in the classical sense in (w_0, w_1) . Let $w \in (w_0, w_1)$. By mean value theorem,

$$(3.26) \quad \frac{f(w) - f(w_0)}{w - w_0} = f'(z) = (r - \lambda)\hat{v}'(z) + (rz - c)\hat{v}''(z) - \lambda\hat{v}'(z) \ln \hat{v}(z)$$

for some $z \in (w_0, w)$. Let $w \rightarrow w_0+$. Notice that $\hat{v}''(z) \rightarrow \infty$ because $\hat{v}''(z) = g(z)$ from equation (3.24), and $g(z)$ has a strictly positive numerator and a denominator that is going to zero from the positive side. So the middle term on the right hand side of (3.26) is exploding to $-\infty$ while the other two terms converge to finite numbers. This contradicts the non-negativity of the left hand side. So $f(w_1) > 0$ necessarily implies $f(w) > 0$ for all $w \in [b, w_1)$. Since $f > 0$ a.e., we conclude that $f > 0$ in $[b, w_s)$. Combining step 1 and 2, we have $\hat{v} \in C^2[b, w_s)$.

From the proof of Lemma 3.5.4, we know $\hat{v}'' > 0$ in (b, w_s) . Optimizing over π in (3.22a) leads to (3.23). Since $\hat{v}'(b) < 0$, (3.23) implies $\hat{v}''(b) > 0$. \square

Once we have C^2 -regularity, we can further upgrade to infinite differentiability

with little effort.

Corollary 3.5.7. $\hat{v} \in C^\infty[b, w_s)$.

Proof. Let g be defined as before. With $\hat{v} \in C^2[b, w_s)$, we now have $g \in C^1[b, w_s)$. It then follows from $\hat{v}'' = g$ that $\hat{v} \in C^3[b, w_s)$. This in turn implies $g \in C^2[b, w_s)$ and so on. Inductively, we will get $\hat{v} \in C^\infty[b, w_s)$. \square

Remark 3.5.8. Since $f(w_s) = 0$, only C^1 -regularity is guaranteed at the right boundary. Even in the non-robust case, it is possible to have an unbounded second derivative at the safe level.

Going back to the original problem through $\hat{u} = \frac{1}{\varepsilon} \ln \hat{v}$, we have the following result.

Proposition 3.5.9. $\hat{u} \in C^1[b, w_s] \cap C^2[b, w_s)$, and satisfies $\hat{u}' < 0$ and $\varepsilon(\hat{u}')^2 + \hat{u}'' > 0$ in $[b, w_s)$. In addition, \hat{u} solves the second order equation

$$(3.27) \quad \lambda u = -R \frac{(u')^2}{\varepsilon(u')^2 + u''} + (rw - c)u', \quad w \in (b, w_s).$$

3.6 Verification

3.6.1 Regularity of π^*

In order to relate \hat{u} to the value function through verification, we first need to show the feedback forms lead to a pair of admissible controls under which the SDE for the controlled wealth process has a unique strong solution. The π attaining the infimum in $F(w, \hat{u}, \hat{u}', \hat{u}'')$ is given by

$$\pi^* = -\frac{\mu - r}{\sigma^2} \frac{\hat{u}'}{\varepsilon(\hat{u}')^2 + \hat{u}''} = -\frac{\mu - r}{\sigma^2} \frac{\hat{v}'}{\hat{v}''},$$

which is the same as the π attaining the infimum in $G(w, \hat{v}, \hat{v}', \hat{v}'')$. We already know from the previous section that π^* is smooth in (b, w_s) , thus locally Lipschitz. We will show π^* is also well-behaved near the boundary.

Lemma 3.6.1.

$$0 < \pi^*(w) < \frac{2(c - rw)}{\mu - r}, \quad w \in [b, w_s).$$

Proof. The lower bound is trivial. For the upper bound, rewrite equation (3.27) as

$$(3.28) \quad \lambda \hat{u} = \left(\frac{\mu - r}{2} \pi^* + rw - c \right) \hat{u}'.$$

For $w \in [b, w_s)$, since $\hat{u}(w) > 0$ and $\hat{u}'(w) < 0$, we must have $\frac{\mu - r}{2} \pi^*(w) + rw - c < 0$. □

Corollary 3.6.2. $\theta^* := \sigma \varepsilon \pi^* \hat{u}'$ is bounded and satisfies $\lim_{w \rightarrow w_s^-} \theta^*(w) = 0$. More precisely,

$$\frac{2\sigma\varepsilon}{\mu - r} (c - rb) \hat{u}'(b) \leq \frac{2\sigma\varepsilon}{\mu - r} (c - rw) \hat{u}'(w) < \theta^*(w) < 0, \quad w \in [b, w_s).$$

It will be verified later that $\pi^*(w)$, $\theta^*(w)$ are the optimal controls for $w \in (b, w_s)$. Observe that the upper bound for π^* given by Lemma 3.6.1 is ϖ . In fact, we can tighten the bound to π_0 and show π^* is non-increasing with respect to ε .

Proposition 3.6.3. $\pi^*(w; \varepsilon)$ is non-increasing in ε for $\varepsilon \geq 0$. In particular, $\pi^*(w; \varepsilon) \leq \pi_0(w)$.

Proof. Let $0 < \varepsilon_1 < \varepsilon_2$ and write $\hat{u}_i(w)$ for $\hat{u}(w; \varepsilon_i)$, $i = 1, 2$. \hat{v}_i and π_i^* are defined similarly. First of all, by comparison principle for problem (3.22), we have $\hat{v}_1 \leq \hat{v}_2$ and thus $\hat{u}_1 \leq \hat{u}_2$. Since $\hat{u}_1(b) = \hat{u}_2(b) = 1$, we deduce

$$\hat{u}'_1(b) = \lim_{w \rightarrow b^+} \frac{\hat{u}_1(w) - 1}{w - b} \leq \lim_{w \rightarrow b^+} \frac{\hat{u}_2(w) - 1}{w - b} = \hat{u}'_2(b).$$

Let $w \rightarrow b^+$ in equation (3.28), we see that

$$\lambda = \left(\frac{\mu - r}{2} \pi^*(b; \varepsilon) + rb - c \right) \hat{u}'(b; \varepsilon).$$

Since $\hat{u}'_1(b) \leq \hat{u}'_2(b) < 0$, we must have $\pi_1^*(b) \geq \pi_2^*(b)$. By Lemma 3.6.1, we also have $\pi_1^*(w_s) = \pi_2^*(w_s) = 0$.⁹ Claim that $\pi_1^*(w) \geq \pi_2^*(w)$ for all $w \in [b, w_s]$.

From equation (3.23), we obtain

$$\lambda \hat{v} \ln \hat{v} = \frac{\mu - r}{2} \hat{v}' \pi^* + (rw - c) \hat{v}', \quad w \in (b, w_s).$$

By Corollary 3.5.7, we can differentiate the above equation. After rearranging terms, we get

$$(3.29) \quad \frac{\mu - r}{2} (\pi^*)' = R + \lambda - r + \frac{\mu - r}{\sigma^2} \frac{rw - c}{\pi^*} + \lambda \ln \hat{v}, \quad w \in (b, w_s).$$

Suppose on the contrary, $\pi_1^* - \pi_2^*$ attains negative minimum at a point $w_0 \in (b, w_s)$.

By first order condition, we have $(\pi_1^*)'(w_0) = (\pi_2^*)'(w_0)$. Equation (3.29) then yields the contradiction:

$$0 = \frac{\mu - r}{\sigma^2} (rw_0 - c) \frac{\pi_1^*(w_0) - \pi_2^*(w_0)}{\pi_1^*(w_0)\pi_2^*(w_0)} + \lambda (\ln \hat{v}_2(w_0) - \ln \hat{v}_1(w_0)) > 0,$$

where we used $\pi_i^* > 0$, $\hat{v}_2 \geq \hat{v}_1$ in (b, w_s) , and the assumption $\pi_1^*(w_0) - \pi_2^*(w_0) < 0$.

Therefore, the claim holds.

If $\varepsilon_1 = 0$, i.e. $\pi_1^* = \pi_0$, then simple computation shows π_1^* satisfies (3.29) with $\hat{v}_1 := 1$. Exactly the same comparison argument implies $\pi_1^* \geq \pi_2^*$ everywhere on $[b, w_s]$. \square

Proposition 3.6.4. π^* is Lipschitz continuous in (b, w_s) and satisfies

$$(3.30) \quad \lim_{w \rightarrow w_s^-} (\pi^*)'(w) = -\frac{\mu - r}{\sigma^2(d - 1)} = \pi_0'(w_s).$$

Proof. For Lipschitz continuity, it suffices to show π^* has bounded first derivative in (b, w_s) . Since $\pi^* > 0$ in $[b, w_s)$, equation (3.29) implies $(\pi^*)'$ is bounded on

⁹By $\pi^*(w_s)$, we mean $\lim_{w \rightarrow w_s^-} \pi^*(w)$ since $\hat{v}''(w)$ may not exist at $w = w_s$.

any subset of (b, w_s) that is away from w_s . It remains to show (3.30). Let $\ell := \liminf_{w \rightarrow w_s^-} (\pi^*)'(w)$ and $L := \limsup_{w \rightarrow w_s^-} (\pi^*)'(w)$. By Proposition 3.6.3, we have

$$\frac{rw - c}{\pi^*} \leq \frac{rw - c}{\pi_0} = -\frac{\sigma^2 r(d-1)}{\mu - r}.$$

The above inequality and (3.29) imply

$$\frac{\mu - r}{2} (\pi^*)' \leq R + \lambda - rd + \lambda \ln \hat{v}.$$

Simple algebra shows $R + \lambda - rd = R/(1-d) < 0$. Since $\hat{v}(w) \rightarrow 1$ as $w \rightarrow w_s^-$, we know $(\pi^*)'$ is negative and bounded away from zero near w_s . In particular, the limit superior L satisfies

$$\frac{\mu - r}{2} L \leq R + \lambda - rd = \frac{R}{1-d} < 0.$$

Now, apply generalized l'Hôpital's rule [78, Theorem II] to (3.29). We deduce

$$\frac{\mu - r}{2} \ell \geq R + \lambda - r + \frac{\mu - r}{\sigma^2} \liminf_{w \rightarrow w_s^-} \frac{r}{(\pi^*)'(w)} = R + \lambda - r + \frac{\mu - r}{\sigma^2} \frac{r}{L}.$$

This leads to a chain of inequalities which is in fact a chain of equalities:

$$0 > \frac{R}{1-d} \geq \frac{\mu - r}{2} L \geq \frac{\mu - r}{2} \ell \geq R + \lambda - r + r(1-d) = \frac{R}{1-d}.$$

So we have proved $\ell = L = \frac{2R}{(\mu-r)(1-d)} = -\frac{\mu-r}{\sigma^2(d-1)}$. □

3.6.2 Verification theorem

For any C^2 function φ and $\pi, \theta \in \mathbb{R}$, define

$$\mathcal{L}^{\pi, \theta} \varphi(w) := [rw - c + (\mu + \sigma\theta - r)\pi] \varphi' + \frac{1}{2} \sigma^2 \pi^2 \varphi''(w).$$

Theorem 3.6.5 (Verification theorem). *Suppose $u : [b, \infty) \rightarrow [0, 1]$, $\Pi : [b, \infty) \rightarrow \mathbb{R}$ and $\Theta : \mathbb{R} \times [b, \infty) \rightarrow \mathbb{R}$ are measurable functions satisfying the following conditions:*

- (i) $u \in C^1[b, w_s] \cap C^2[b, w_s)$;

(ii) u is a solution of

$$(3.31) \quad \lambda u(w) = \inf_{\pi} \sup_{\theta} \left\{ -\frac{1}{2\varepsilon} \theta^2 + \mathcal{L}^{\pi, \theta} u(w) \right\}, \quad w \in (b, w_s);$$

(iii) $u(b) = 1$ and $u(w) = 0$ for $w \geq w_s$;

(iv) $\Pi(w)$ attains the infimum in (ii) for each $w \in (b, w_s)$; $\Theta(\pi, w)$ attains the supremum in (ii) for each $\pi \in \mathbb{R}$ and $w \in (b, w_s)$;

(v) $\Pi(w) = \Theta(\pi, w) = 0$ if $w \notin (b, w_s)$;

(vi) Π is bounded and Lipschitz continuous in (b, w_s) ; Θ is bounded on $[\pi_1, \pi_2] \times [b, \infty)$ for any compact interval $[\pi_1, \pi_2] \subset \mathbb{R}$.

Then $\psi = u$ on $[b, \infty)$, and $\Pi(\cdot), \Theta(\Pi(\cdot), \cdot)$ are optimal Markovian controls.

Proof. Same as [14], we let Δ be the ‘‘coffin state’’ and $[b, \infty) \cup \{\Delta\}$ be the one point compactification of $[b, \infty)$. Define the extension of u to $[b, \infty) \cup \{\Delta\}$ by assigning $u(\Delta) = 0$.

1. Let $w > b$. By conditions (v) and (vi), the SDE

$$dW_t = [rW_t + (\mu - r)\Pi(W_t) - c]dt + \sigma\Pi(W_t)dB_t, \quad W_0 = w$$

has a unique strong solution $W^{w, \Pi}$ w.r.t. the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Let $\pi_t^* := \Pi(W_t^{w, \Pi})$ and write $W^{w, \pi^*} := W^{w, \Pi}$. $\pi^* \in \mathcal{A}$ since Π is bounded and measurable. Define $\tau_b^* := \inf\{t \geq 0 : W^{w, \pi^*} \leq b\}$ and $\tau^* := \inf\{t \geq 0 : W^{w, \pi^*} \geq w_s\} \wedge \tau_d$.

Let $\mathbb{Q} \in \mathcal{Q}$ be any candidate measure with corresponding drift distortion process θ . By Girsanov theorem, $B_t^{\mathbb{Q}} := B_t - \int_0^t \theta_s ds$ is a \mathbb{Q} -Brownian motion. W^{w, π^*}

satisfies

$$dW_t = [rW_t + (\mu + \sigma\theta_t - r)\pi_t^* - c]dt + \sigma\pi_t^*dB_t^{\mathbb{Q}}, \quad W_0 = w.$$

Recall that τ_d is the first jump time of the \mathbb{P} -Poisson process N with rate λ that is independent of \mathbb{F} . The definition of \mathcal{Q} ensures that N is also a \mathbb{Q} -Poisson process with the same rate. Let $\overline{W}_t^{w, \pi^*} := W_t^{w, \pi^*} 1_{\{t < \tau_d\}} + \Delta 1_{\{t \geq \tau_d\}}$. \overline{W}^{w, π^*} is a progressively measurable process in the enlarged filtration \mathbb{H} which includes information generated by N . It is easy to see that \overline{W}^{w, π^*} satisfies:

$$dW_t = [rW_t + (\mu + \sigma\theta_t - r)\pi_t^* - c]dt + \sigma\pi_t^*dB_t^{\mathbb{Q}} - (\Delta - W_{t-})dN_t, \quad W_0 = w.$$

Applying Itô's lemma to $u(\overline{W}_t^{w, \pi^*})$ and using that $u(\Delta) = 0$, we have

$$\begin{aligned} u(\overline{W}_{\tau_b^* \wedge \tau^*}^{w, \pi^*}) &= u(w) + \int_0^{\tau_b^* \wedge \tau^*} \mathcal{L}^{\pi_s^*, \theta_s} u(W_s^{w, \pi^*}) - \lambda u(W_s^{w, \pi^*}) ds \\ &\quad + \int_0^{\tau_b^* \wedge \tau^*} u'(W_s^{w, \pi^*}) \sigma \pi_s^* dB_s^{\mathbb{Q}} - u(W_{s-}^{w, \pi^*}) d(N_s - \lambda s). \end{aligned}$$

Since u , u' and Π are bounded on $[b, w_s]$, the Itô integral vanishes upon taking \mathbb{Q} -expectation and we get

$$\mathbb{E}^{\mathbb{Q}} \left[u(\overline{W}_{\tau_b^* \wedge \tau^*}^{w, \pi^*}) \right] = u(w) + \mathbb{E}^{\mathbb{Q}} \left[\int_0^{\tau_b^* \wedge \tau^*} \mathcal{L}^{\pi_s^*, \theta_s} u(W_s^{w, \pi^*}) - \lambda u(W_s^{w, \pi^*}) ds \right].$$

Conditions (ii), (iv) and that $\pi^* = \Pi(W^{w, \pi^*})$ imply for $0 \leq s < \tau_b^* \wedge \tau^*$,

$$\begin{aligned} 0 &= \inf_{\pi} \sup_{\theta} \left\{ -\frac{1}{2\varepsilon} \theta^2 + \mathcal{L}^{\pi, \theta} u(W_s^{w, \pi^*}) \right\} - \lambda u(W_s^{w, \pi^*}) \\ &= \sup_{\theta} \left\{ -\frac{1}{2\varepsilon} \theta^2 + \mathcal{L}^{\pi_s^*, \theta} u(W_s^{w, \pi^*}) \right\} - \lambda u(W_s^{w, \pi^*}) \\ &\geq -\frac{1}{2\varepsilon} \theta_s^2 + \mathcal{L}^{\pi_s^*, \theta_s} u(W_s^{w, \pi^*}) - \lambda u(W_s^{w, \pi^*}). \end{aligned}$$

So we have

$$\mathbb{E}^{\mathbb{Q}} \left[u(\overline{W}_{\tau_b^* \wedge \tau^*}^{w, \pi^*}) \right] \leq u(w) + \mathbb{E}^{\mathbb{Q}} \left[\int_0^{\tau_b^* \wedge \tau^*} \frac{1}{2\varepsilon} \theta_s^2 ds \right].$$

Equivalently,

$$u(w) \geq \mathbb{E}^{\mathbb{Q}} \left[u(\overline{W}_{\tau_b^* \wedge \tau^*}^{w, \pi^*}) - \int_0^{\tau_b^* \wedge \tau^*} \frac{1}{2\varepsilon} \boldsymbol{\theta}_s^2 ds \right].$$

By condition (v), W^{w, π^*} will stay constant once it reaches the safe level. This means, if the safe level is reached, then death will definitely occur before ruin. So we have $\{\tau_b^* < \tau^*\} = \{\tau_b^* < \tau_d^*\}$. Since $u(\overline{W}_{\tau_b^* \wedge \tau^*}^{w, \pi^*}) = 1_{\{\tau_b^* < \tau^*\}} = 1_{\{\tau_b^* < \tau_d^*\}}$ and $\tau_b^* \wedge \tau^* \leq \tau_d$, we get

$$u(w) \geq \mathbb{E}^{\mathbb{Q}} \left[1_{\{\tau_b^* < \tau_d\}} - \int_0^{\tau_d} \frac{1}{2\varepsilon} \boldsymbol{\theta}_s^2 ds \right].$$

This holds for all $\mathbb{Q} \in \mathcal{Q}$. So

$$\begin{aligned} u(w) &\geq \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{\tau_b^* < \tau_d\}} - \int_0^{\tau_d} \frac{1}{2\varepsilon} \boldsymbol{\theta}_s^2 ds \right] \\ &\geq \inf_{\boldsymbol{\pi} \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{\tau_b^{w, \boldsymbol{\pi}} < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_d} \boldsymbol{\theta}_s^2 ds \right] = \psi(w), \end{aligned}$$

where we put superscripts on τ_b in the last step to indicate its dependence on the initial wealth and the control.

2. Let $\boldsymbol{\pi} \in \mathcal{A}$ be any admissible investment strategy and $W^{w, \boldsymbol{\pi}}$ be the solution to the SDE:

$$dW_t = [rW_t + (\mu - r)\boldsymbol{\pi}_t - c]dt + \sigma\boldsymbol{\pi}_t dB_t, \quad W_0 = w.$$

Let $\boldsymbol{\theta}_t^* := \Theta(\boldsymbol{\pi}_t, W_t^{w, \boldsymbol{\pi}})$. $\boldsymbol{\theta}^*$ is \mathbb{F} -progressively measurable since both $\boldsymbol{\pi}$ and $W^{w, \boldsymbol{\pi}}$ are, and Θ is a measurable function. Since $\boldsymbol{\pi}_t$ is a.s. bounded uniformly in t , condition (vi) ensures $\boldsymbol{\theta}^*$ satisfies all integrability conditions in the definition of \mathcal{Q} . So there exists a measure $\mathbb{Q}^* \in \mathcal{Q}$ satisfying $\frac{d\mathbb{Q}_t^*}{d\mathbb{P}_t} = \mathcal{E}(\int_0^t \boldsymbol{\theta}_s^* dB_s)$ where \mathcal{E} denotes the stochastic exponential. It follows from Girsanov theorem that $B_t^{\mathbb{Q}^*} := B_t - \int_0^t \boldsymbol{\theta}_s^* ds$ is a \mathbb{Q}^* -Brownian motion. So $W^{w, \boldsymbol{\pi}}$ satisfies:

$$dW_t = [rW_t + (\mu + \sigma\boldsymbol{\theta}_t^* - r)\boldsymbol{\pi}_t - c]dt + \sigma\boldsymbol{\pi}_t dB_t^{\mathbb{Q}^*}, \quad W_0 = w.$$

Define $\tau_b := \inf\{t \geq 0 : W_t^{w,\pi} \leq b\}$ and $\tau := \inf\{t \geq 0 : W_t^{w,\pi} \geq w_s\} \wedge \tau_d$. Same as before, we work with the larger filtration \mathbb{H} and consider the process $\overline{W}^{w,\pi} = W_t^{w,\pi} 1_{\{t < \tau_d\}} + \Delta 1_{\{t \geq \tau_d\}}$ which satisfies the SDE:

$$dW_t = [rW_t + (\mu + \sigma\theta_t^* - r)\pi_t - c]dt + \sigma\pi_t dB_t^{\mathbb{Q}^*} + (\Delta - W_{t-})dN_t, \quad W_0 = w.$$

Again, thanks to the drift distortion θ^* being \mathbb{F} -adapted, N remains a Poisson process with rate λ under \mathbb{Q}^* . By Itô's lemma and that $u(\Delta) = 0$, we have for any $t \geq 0$,

$$\begin{aligned} u(\overline{W}_{\tau_b \wedge \tau \wedge t}^{w,\pi}) &= u(w) + \int_0^{\tau_b \wedge \tau \wedge t} -\lambda u(W_s^{w,\pi}) + \mathcal{L}^{\pi_s, \theta_s^*} u(W_s^{w,\pi}) ds \\ &\quad + \int_0^{\tau_b \wedge \tau \wedge t} u'(W_s^{w,\pi}) \sigma \pi_s dB_s^{\mathbb{Q}^*} - u(W_{s-}^{w,\pi}) d(N_s - \lambda s). \end{aligned}$$

Taking \mathbb{Q}^* expectation yields

$$\mathbb{E}^{\mathbb{Q}^*} [u(\overline{W}_{\tau_b \wedge \tau \wedge t}^{w,\pi})] = u(w) + \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^{\tau_b \wedge \tau \wedge t} -\lambda u(W_s^{w,\pi}) + \mathcal{L}^{\pi_s, \theta_s^*} u(W_s^{w,\pi}) ds \right],$$

where the Itô integral vanishes because u, u' are bounded and π is \mathbb{Q}_t^* -a.s. bounded for all $t \geq 0$. By conditions (ii), (iv), and our definition of θ^* , we know

$$\begin{aligned} (3.32) \quad 0 &= \inf_{\pi} \sup_{\theta} \left\{ -\frac{1}{2\varepsilon} \theta^2 + \mathcal{L}^{\pi, \theta} u(W_s^{w,\pi}) \right\} - \lambda u(W_s^{w,\pi}) \\ &\leq \sup_{\theta} \left\{ -\frac{1}{2\varepsilon} \theta^2 + \mathcal{L}^{\pi_s, \theta} u(W_s^{w,\pi}) \right\} - \lambda u(W_s^{w,\pi}) \\ &= -\frac{1}{2\varepsilon} (\theta_s^*)^2 + \mathcal{L}^{\pi_s, \theta_s^*} u(W_s^{w,\pi}) - \lambda u(W_s^{w,\pi}) \end{aligned}$$

for $s \in [0, \tau_b \wedge \tau)$. So

$$(3.33) \quad \mathbb{E}^{\mathbb{Q}^*} [u(\overline{W}_{\tau_b \wedge \tau \wedge t}^{w,\pi})] \geq u(w) + \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^{\tau_b \wedge \tau \wedge t} \frac{1}{2\varepsilon} (\theta_s^*)^2 ds \right].$$

Letting $t \rightarrow \infty$ and using bounded and monotone convergence theorems, we get

$$(3.34) \quad \mathbb{E}^{\mathbb{Q}^*} [u(\overline{W}_{\tau_b \wedge \tau}^{w,\pi})] \geq u(w) + \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^{\tau_b \wedge \tau} \frac{1}{2\varepsilon} (\theta_s^*)^2 ds \right].$$

Since $u(\overline{W}_{\tau_b \wedge \tau}^{w, \pi}) = 1_{\{\tau_b < \tau\}} \leq 1_{\{\tau_b < \tau_d\}}$, we obtain

$$(3.35) \quad u(w) \leq \mathbb{E}^{\mathbb{Q}^*} \left[1_{\{\tau_b < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_b \wedge \tau} (\boldsymbol{\theta}_s^*)^2 ds \right].$$

Let us first assume $\boldsymbol{\pi}$ is an admissible strategy such that $\boldsymbol{\pi} = 0$ once ruin occurs or safe level is reached, so that the wealth process will stay at the ruin or safe level until death time. Denote by \mathcal{A}_0 the collection of such subclass of strategies. Then by condition (v), we have $\boldsymbol{\theta}_s^* = 0$ for $\tau_b \wedge \tau < s < \tau_d$. Hence

$$u(w) \leq \mathbb{E}^{\mathbb{Q}^*} \left[1_{\{\tau_b < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_d} \boldsymbol{\theta}_s^* ds \right] \leq \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{\tau_b < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_d} \boldsymbol{\theta}_s^2 ds \right].$$

This holds for any $\boldsymbol{\pi} \in \mathcal{A}_0$. So we have

$$u(w) \leq \inf_{\boldsymbol{\pi} \in \mathcal{A}_0} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{\tau_b^{w, \boldsymbol{\pi}} < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_d} \boldsymbol{\theta}_s^2 ds \right],$$

where we added superscripts to τ_b to indicate its dependence on the initial wealth and the control. It remains to note that controls in $\mathcal{A} \setminus \mathcal{A}_0$ do not yield a smaller infimum because once ruin occurs, it becomes a history that cannot be altered; once safe level is reached, no policy can do a better job than zero ruin probability. Therefore, we actually have

$$u(w) \leq \inf_{\boldsymbol{\pi} \in \mathcal{A}} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}} \left[1_{\{\tau_b^{w, \boldsymbol{\pi}} < \tau_d\}} - \frac{1}{2\varepsilon} \int_0^{\tau_d} \boldsymbol{\theta}_s^2 ds \right] = \psi(w).$$

3. As in step 1, let $W^{w, \Pi}$ be the unique strong solution of

$$dW_t = [rW_t + (\mu - r)\Pi(W_t) - c]dt + \sigma\Pi(W_t)dB_t, \quad W_0 = w,$$

and $\boldsymbol{\pi}^* := \Pi(W^{w, \Pi}) \in \mathcal{A}$. Let $\boldsymbol{\theta}_t^* := \Theta(\Pi(W_t^{w, \Pi}), W_t^{w, \Pi})$. Conditions (v), (vi) and the measurability of Θ, Π ensures $\boldsymbol{\theta}^*$ is a bounded, \mathbb{F} -progressively measurable process. So there is a measure $\mathbb{Q}^* \in \mathcal{Q}$ having $\boldsymbol{\theta}^*$ as the corresponding drift distortion process. Repeat the analysis in step 2 using controls $\boldsymbol{\pi}^*$ and $\boldsymbol{\theta}^*$. (3.32) through

(3.34) now hold with equality. For (3.35), since π^* and θ^* will both be zero and remain zero until death time once the ruin level or the safe level is reached, we have

$$\{\tau_b^* < \tau^*\} = \{\tau_b^* < \tau_d\} \text{ and}$$

$$\psi(w) = u(w) = \mathbb{E}^{\mathbb{Q}^*} \left[1_{\{\tau_b^* < \tau_d\}} - \int_0^{\tau_d} \frac{1}{2\varepsilon} (\theta_s^*)^2 ds \right],$$

where τ_b^* and τ^* denote the ruin time and the minimum of safe and death times, respectively, when the wealth starts at w and is controlled by π^* . This proves the optimality of the feedback forms. \square

3.6.3 Proof of Theorem 3.2.4.

Proof. The functions

$$u(w) := \hat{u}(w)1_{\{w \leq w_s\}}, \quad \Pi(w) := \pi^*(w)1_{(b, w_s)}(w), \quad \Theta(\pi, w) := \sigma \varepsilon \pi \hat{u}'(w)1_{(b, w_s)}(w)$$

satisfy all conditions of the verification theorem. (i) follows from Proposition 3.5.9. (ii) and (iii) hold because \hat{u} solves (3.15) and $F = 0$ is equivalent to (3.31). (iv) follows from first order conditions and the definition of π^* (see the beginning of Section 3.6). (v) is clear from the definition of Π and Θ . (vi) holds by Propositions 3.6.4 and 3.5.9; the latter implies \hat{u}' is bounded on $[b, w_s]$. \square

Remark 3.6.6. Verification can be carried out even if π^* is only known to be locally Lipschitz continuous, because the optimally controlled wealth process actually never reaches the safe level (see Proposition 3.7.3). What Proposition 3.6.4 shows on top of the global Lipschitz continuity is that $\pi^*(w; \varepsilon)$ is tangent to $\pi_0(w)$ at $w = w_s$ for all $0 < \varepsilon < \infty$.

In the remaining sections of this chapter, we will speak of π^* , θ^* given by (3.10) and (3.11) as the optimal Markovian controls. It is understood that they are optimal in the interval (b, w_s) .

3.7 Properties of the value function and the optimal investment policy

Let us first summarize some properties of ψ and π^* that we have already seen.

- (i) $\psi \in C^1[b, w_s] \cap C^2(b, w_s)$ and is strictly decreasing on $[b, w_s]$;
- (ii) ψ is non-decreasing in ε , bounded from below by ψ_0 and from above by $\psi_\infty \wedge \mathfrak{p}$;
- (iii) $0 < \pi^* \leq \pi_0$ in $[b, w_s)$ and π^* is non-increasing in ε ;
- (iv) π^* is Lipschitz continuous in (b, w_s) and is tangent to π_0 at the safe level.

In this section, we prove two additional properties. The first one reveals how the concavity of ψ depends on parameters. The second one addresses the question of whether the safe level can be reached by the optimally controlled wealth process. In the non-robust case, [82] shows it is never reached in finite time. Same phenomenon exists for our robust problem; the individual either loses the game, or “wins” the game by dying.

3.7.1 When does the value function lose convexity?

Proposition 3.7.1.

- (i) If $r \leq \lambda$, then ψ is convex on $[b, w_s]$. If $r < \lambda$, ψ is strictly convex.
- (ii) If $r > \lambda$, then ψ changes concavity at most once on $[b, w_s]$. If $0 \leq \varepsilon \leq \frac{R}{rd-\lambda}$, ψ is strictly convex on $[b, w_s]$. If $\varepsilon > \frac{R}{r-\lambda}$, ψ is strictly concave in $[b, w_0)$ and strictly convex in $(w_0, w_s]$ where w_0 is the unique point in (b, w_s) satisfying $(rw_0 - c)\psi'(w_0) - \lambda\psi(w_0) = \frac{R}{\varepsilon}$.

Proof. (i) When $\varepsilon = 0$, strict convexity holds regardless of the sign of $r - \lambda$. Assume $\varepsilon > 0$. Let $f(w) := (rw - c)\psi' - \lambda\psi$. When proving Proposition 3.5.6, we showed

that $(rw - c)\hat{v}' - \lambda\hat{v} \ln \hat{v} > 0$ in $[b, w_s]$. In terms of ψ which equals \hat{u} on $[b, w_s]$, we have $\varepsilon e^{\varepsilon\psi}[(rw - c)\psi' - \lambda\psi] > 0$, which implies $f > 0$ on $[b, w_s]$. Recall that ψ satisfies

$$R \frac{(\psi')^2}{\varepsilon(\psi')^2 + \psi''} = f.$$

Moving ψ'' to one side and everything else to the other side, we obtain

$$(3.36) \quad \psi'' = \left(\frac{R}{f} - \varepsilon \right) (\psi')^2.$$

We see that the sign of ψ'' depends on the relative size of f to R/ε . Since $\psi \in C^2[b, w_s]$, we can differentiate f and get

$$(3.37) \quad f' = (r - \lambda)\psi' + (rw - c)\psi'' \geq (rw - c)\psi'',$$

where the inequality follows from $\psi' < 0$ and the assumption that $r \leq \lambda$. Since $f(w_s) = 0$, f attains maximum either at an interior point or at $w = b$. In both cases, we have $f'(w_m) \leq 0$ where $w_m \in [b, w_s]$ is the point where maximum is attained. It follows from (3.37) that $\psi''(w_m) \geq 0$ and then from (3.36) that $f(w_m) \leq R/\varepsilon$. Since w_m is a maximum point, $f(w) \leq R/\varepsilon$ for all $w \in [b, w_s]$. This in turn implies by (3.36) that $\psi''(w) \geq 0$ for all $w \in (b, w_s)$. Since ψ is continuous, interior convexity can be extended to the boundary. If $r < \lambda$, then the inequality in (3.37) becomes strict and we have $\psi''(w_m) > 0$ at the point w_m of maximality of f . Subsequent inequalities all become strict and we obtain strict convexity of ψ .

(ii) First of all, equation (3.36) implies in any circumstances, regardless of the sign of $r - \lambda$ and the value of ε , ψ will be strictly convex in a neighborhood of w_s . This is because $f(w_s) = 0$ so that $R/f(w) - \varepsilon > 0$ for w sufficiently close to w_s . Let $r > \lambda$. Then (3.37) becomes

$$(3.38) \quad f' = (r - \lambda)\psi' + (rw - c)\psi'' < (rw - c)\psi'', \quad w \in [b, w_s].$$

If ψ changes concavity at w_0 , then $\psi''(w_0) = 0$ and the above inequality implies $f'(w_0) < 0$. So f is strictly decreasing whenever ψ changes concavity.¹⁰ Looking at (3.36), we deduce that ψ can only change from concave to convex if concavity changes at all. Since we have already argued ψ is strictly convex in a neighborhood of w_s , we conclude that if ψ is not convex everywhere, then it changes concavity only once; it is strictly concave up to the (unique) point w_0 where $f(w_0) = R/\varepsilon$ and is strictly convex afterwards. We also note that since f can only touch or cross the horizontal line at R/ε in a decreasing fashion, $f(w) > R/\varepsilon$ for $w \in [b, w_0)$ and $f(w) < R/\varepsilon$ for $w \in (w_0, w_s]$.

Next, we identify some cases when ψ changes or does not change concavity. In view of the way f intersect the horizontal line at R/ε , it suffices to check whether $f(b) > R/\varepsilon$. We have

$$f(b) = (rb - c)\psi'(b) - \lambda.$$

If $0 \leq \varepsilon \leq \frac{R}{rd - \lambda}$, then $f(b) \leq (rb - c)\psi'_0(b) - \lambda = rd - \lambda \leq R/\varepsilon$ and there will be no concavity change for ψ .

If $\varepsilon > \frac{R}{r - \lambda}$, then we consider two cases. If $\psi'(b) > \frac{1}{b - w_s} = \frac{r}{rb - c}$, then ψ cannot be convex everywhere. If it is convex everywhere, it will stay above its tangent line passing through the point $(b, 1)$. But the point $(w_s, 0)$ lies below this tangent line, which means the right boundary condition is not satisfied. If $\psi'(b) \leq \frac{r}{rb - c}$, then $f(b) \geq r - \lambda > R/\varepsilon$. In both cases, ψ will change concavity. \square

Remark 3.7.2. Based on the feedback form (3.11), θ^* is bounded below by $-\frac{\mu - r}{\sigma}$ whenever ψ is convex. If ψ is not convex, then its infection point is the unique point where $\theta^* = -\frac{\mu - r}{\sigma}$. In other words, ψ changes concavity when the distorted

¹⁰ ψ cannot be locally linear because otherwise on one hand, (3.38) implies $f' < 0$ and f is locally strictly decreasing; on the other hand, (3.36) implies f is locally constant.

Sharpe ratio $\frac{\mu-r}{\sigma} + \theta^*$ is zero. Moreover, since ψ changes from concave to convex, the distorted Sharpe ratio is negative to the left of this inflection point and positive to the right of this inflection point.

3.7.2 Reaching the safe level or not?

Proposition 3.7.3. *Let $b < w < w_s$ and W^* be the optimally controlled wealth starting at w . Let $\tau_s^* := \inf\{t \geq 0 : W_t^* \geq w_s\}$ and $\tau_b^* := \inf\{t \geq 0 : W_t^* \leq b\}$. Then $\mathbb{P}(\tau_s^* < \tau_b^*) = 0$.*

Proof. Since we are only interested in whether the safe level can be reached before ruin, we may extend the domain of Π to \mathbb{R} and set $\Pi(w) := \frac{c-rw}{\mu-r}$ for $w \leq b$. Let \widetilde{W} be the solution to the SDE:

$$dW_t = [rW_t + (\mu - r)\Pi(W_t) - c]dt + \sigma\Pi(W_t)dB_t, \quad W_0 = w.$$

\widetilde{W} equals W^* up to ruin time. It suffices to show \widetilde{W} does not exit the interval $(-\infty, w_s)$ in finite time, and we use Feller's test for explosions (see section 5.5.C of [57]). By Lemma 3.6.1, non-degeneracy and local integrability hold on this interval.

Let $\mathfrak{s}(w) := \sigma\Pi(w)$ and $\mathfrak{b}(w) := rw + (\mu - r)\Pi(w) - c$. Fix $w_0 \in (-\infty, w_s)$. Let

$$p(w) := \int_{w_0}^w \exp\left(-2 \int_{w_0}^y \frac{\mathfrak{b}(z)}{\mathfrak{s}^2(z)} dz\right) dy$$

be the scale function, and

$$v(w) := \int_{w_0}^w p'(y) \int_{w_0}^y \frac{2dz}{p'(z)\mathfrak{s}^2(z)} dy = \int_{w_0}^w \int_{w_0}^y \frac{2}{\mathfrak{s}^2(z)} \exp\left(-2 \int_z^y \frac{\mathfrak{b}(x)}{\mathfrak{s}^2(x)} dx\right) dz dy.$$

We want to show $v(-\infty) = v(w_s) = \infty$. $v(-\infty) = \infty$ is easy by the way we extend

Π . Let $a \leq w_0 \wedge b$. Since $\mathfrak{b}(x) = 0$ for $x \leq a$, we have

$$\begin{aligned} v(-\infty) &= \int_{-\infty}^{w_0} \int_y^{w_0} \frac{2}{\mathfrak{s}^2(z)} \exp\left(2 \int_y^z \frac{\mathfrak{b}(x)}{\mathfrak{s}^2(x)} dx\right) dz dy \geq \int_{-\infty}^a \int_y^a \frac{2}{\mathfrak{s}^2(z)} dz dy \\ &= \int_{-\infty}^a \int_y^a \frac{4R}{(c - rz)^2} dz dy = \infty. \end{aligned}$$

To show $v(w_s) = \infty$, we use Lemma 3.6.1 or Proposition 3.6.3 to obtain $\Pi(w) \leq K_1(c - rw)$, $w \in (b, w_s)$ for some positive constant K_1 . It follows that $|\mathbf{b}(w)| \leq [1 + K_1(\mu - r)](c - rw)$, $w \in (b, w_s)$. Also observe that if $\mathbf{b}(w) > 0$, then $\Pi(w) > \frac{c - rw}{\mu - r}$.

So we have

$$\frac{\mathbf{b}(w)}{\mathfrak{s}^2(w)} \leq 1_{\{\mathbf{b}(w) > 0\}} \frac{\mathbf{b}(w)}{\mathfrak{s}^2(w)} \leq 1_{\{\mathbf{b}(w) > 0\}} \frac{2R[1 + K_1(\mu - r)]}{c - rw} \leq \frac{K_2}{c - rw}, w \in (b, w_s),$$

where $K_2 := 2R[1 + (\mu - r)K_1] > 0$. Let $(b \vee w_0) \leq a' < w_s$.

$$\begin{aligned} v(w_s) &= \int_{w_0}^{w_s} \int_{w_0}^y \frac{2}{\mathfrak{s}^2(z)} \exp\left(-2 \int_z^y \frac{\mathbf{b}(x)}{\mathfrak{s}^2(x)} dx\right) dz dy \\ &\geq \int_{a'}^{w_s} \int_{a'}^y \frac{2}{\sigma^2 K_1^2 (c - rz)^2} \exp\left(-2 \int_z^y \frac{K_2}{c - rx} dx\right) dz dy \\ &= \int_{a'}^{w_s} \int_{a'}^y \frac{2}{\sigma^2 K_1^2 (c - rz)^2} \left(\frac{c - ry}{c - rz}\right)^{\frac{2K_2}{r}} dz dy \\ &= \frac{2}{\sigma^2 K_1^2 (r + 2K_2)} \int_{a'}^{w_s} (c - ry)^{\frac{2K_2}{r}} \left[(c - ry)^{-1 - \frac{2K_2}{r}} - (c - ra')^{-1 - \frac{2K_2}{r}}\right] dy \\ &= \frac{2}{\sigma^2 K_1^2 (r + 2K_2)} \left[\int_{a'}^{w_s} \frac{1}{c - ry} dy - \int_{a'}^{w_s} \frac{1}{c - ra'} \left(\frac{c - ry}{c - ra'}\right)^{\frac{2K_2}{r}} dy \right]. \end{aligned}$$

The second integral is finite while the first integral diverges to ∞ . So we obtain $v(w_s) = \infty$. The rest is by Feller's test for explosions. \square

3.8 Asymptotic expansion for small ε

In general, (3.12) does not have an explicit solution, but it turns out that for small ε , there are explicit formulas for the leading term and the first order correction.

Rewrite (3.12a) as

$$(3.39) \quad ((rw - c)\psi' - \lambda\psi) (\varepsilon(\psi')^2 + \psi'') = R(\psi')^2.$$

Let

$$f_0(w) + f_1(w)\varepsilon + f_2(w)\varepsilon^2 + \dots$$

be an asymptotic expansion of $\psi(w)$ as $\varepsilon \rightarrow 0$. Substituting the expansion into (3.39) and collecting zero-th order terms in ε , we get

$$(3.40) \quad ((rw - c)f'_0 - \lambda f_0) f''_0 = R(f'_0)^2$$

which is precisely the differential equation satisfied by the non-robust value function.

We impose the boundary conditions $f_0(b) = 1$ and $f_0(w_s) = 0$. Then

$$f_0(w) = \psi_0(w) = \left(\frac{c - rw}{c - rb} \right)^d.$$

Collecting first order terms in ε , we get

$$(3.41) \quad [(rw - c)f'_1 - \lambda f_1]f''_0 + [(rw - c)f'_0 - \lambda f_0][(f'_0)^2 + f''_0] = 2Rf'_0f'_1.$$

Using the formula for f_0 , after some computation, we arrive at a linear second order ODE for f_1 :

$$(3.42) \quad f''_1 + A(w)f'_1 + B(w)f_1 + C(w) = 0$$

where

$$\begin{aligned} A(w) &:= \frac{r(d-1)(2R - rd + r)}{R} \frac{1}{c - rw}, \\ B(w) &:= \frac{-\lambda r^2(d-1)^2}{R} \frac{1}{(c - rw)^2}, \\ C(w) &:= \frac{r^2 d^2}{(c - rb)^{2d}} (c - rw)^{2d-2}. \end{aligned}$$

We require f_1 to satisfy the homogeneous boundary conditions $f_1(b) = f_1(w_s) = 0$.

Let $x = c - rw$ and $g(x) = f_1(w)$. Equation (3.42) can be rewritten as

$$(3.43) \quad x^2 g'' - \frac{(d-1)(2R - rd + r)}{R} x g' - \frac{\lambda(d-1)^2}{R} g + \frac{d^2}{(c - rb)^{2d}} x^{2d} = 0$$

with boundary conditions $g(0) = g(c - rb) = 0$. This is a non-homogeneous Cauchy-Euler equation. The corresponding homogeneous equation has general solution:

$$g_h(x) = C_1 x^{k_1} + C_2 x^{k_2}$$

where $k_1 > 0 > k_2$ are the roots of

$$k^2 - \left(2d - 1 - \frac{r(d-1)^2}{R}\right)k - \frac{\lambda(d-1)^2}{R} = 0.$$

It turns out that $k_1 = d$. For a particular solution, we guess the form $g_p(x) = C_p x^{2d}$.

Substituting g_p into (3.43), we find

$$C_p = \frac{-Rd^2}{(c-rb)^{2d} [(d-1)^2(2dr-\lambda) + 2Rd]}$$

The general solution to (3.43) is $g = g_h + g_p$. Since $g(0) = 0$, we must have $C_2 = 0$, otherwise the solution would explode at $x = 0$. The other boundary condition $g(c-rb) = 0$ yields

$$C_1 = -C_p(c-rb)^d.$$

So we have obtained

$$f_1(w) = g(c-rw) = \frac{Rd^2}{(d-1)^2(2dr-\lambda) + 2Rd} \left[\left(\frac{c-rw}{c-rb}\right)^d - \left(\frac{c-rw}{c-rb}\right)^{2d} \right].$$

Proposition 3.8.1.

$$\psi(w) = \psi_0(w) + \frac{Rd^2(\psi_0(w) - \psi_0^2(w))}{(d-1)^2(2dr-\lambda) + 2Rd} \varepsilon + O(\varepsilon^2)$$

as $\varepsilon \downarrow 0$ uniformly in w , where the constants R, d are defined in (3.5).

Proof. We only give a sketch proof. Let $\tilde{\psi} := f_0 + f_1\varepsilon$. We want to show $\psi(w) = \tilde{\psi}(w) + O(\varepsilon^2)$ uniformly for $w \in (b, w_s)$. Using the formulas for f_0 and f_1 , we can show for ε sufficiently small,

$$(3.44) \quad \varepsilon(\tilde{\psi}')^2(w) + \tilde{\psi}''(w) \geq C_1 \left(\frac{c-rw}{c-rb}\right)^{d-2} > 0 \quad \forall w \in (b, w_s).$$

for some positive constant C_1 independent of ε and w . Next, we show

$$(3.45) \quad F(w, \tilde{\psi}(w), \tilde{\psi}'(w), \tilde{\psi}''(w)) = O(\varepsilon^2) \quad \text{uniformly for } w \in (b, w_s).$$

In view of (3.44), we carry out the optimization over π in the expression for F . Using equations (3.40) and (3.41), we obtain

$$F(w, \tilde{\psi}(w), \tilde{\psi}'(w), \tilde{\psi}''(w)) = \frac{D_1(w)\varepsilon^2 + D_2(w)\varepsilon^3 + D_3(w)\varepsilon^4}{\varepsilon(\tilde{\psi}')^2(w) + \tilde{\psi}''(w)},$$

where

$$D_1(w) := 2f_0'f_1'[\lambda f_0 - (rw - c)f_0'] + [\lambda f_1 - (rw - c)f_1'][(f_0')^2 + f_1''] + R(f_1')^2,$$

$$D_2(w) := 2f_0'f_1'[\lambda f_1 - (rw - c)f_1'] + (f_1')^2[\lambda f_0 - (rw - c)f_0'],$$

$$D_3(w) := (f_1')^2[\lambda f_1 - (rw - c)f_1'].$$

It can be shown using (3.44) and the explicit formulas for f_0, f_1 that there exists a positive constant C_2 independent of ε and w , such that $|D_i(w)|/[\varepsilon(\tilde{\psi}')^2(w) + \tilde{\psi}''(w)] \leq C_2, i = 1, 2, 3$ for all $w \in (b, w_s)$ and for ε small enough. This proves (3.45). Consequently, we can find a positive constant C_3 such that for ε sufficiently small,

$$F(w, \tilde{\psi}(w) - C_3\varepsilon^2, \tilde{\psi}'(w), \tilde{\psi}''(w)) \leq 0 \quad \text{and} \quad F(w, \tilde{\psi}(w) + C_3\varepsilon^2, \tilde{\psi}'(w), \tilde{\psi}''(w)) \geq 0.$$

By comparison principle for the equation $F = 0$, we have $\psi = \tilde{\psi} + O(\varepsilon^2)$. \square

3.9 Numerical examples

We solve the boundary value problem (3.12) numerically using finite difference method. The model parameters used are $c = 1, b = 1, \mu = 0.1, \sigma = 0.15, \lambda = 0.04$ and $\varepsilon = 0, 1, 5, 10, 50$. We choose a hazard rate of 0.04, i.e. an expected future lifetime of 25 years, because the investment problem we considered is more relevant to retirees. To demonstrate that the concavity of the value function is closely related to how interest rate compares with hazard rate, we work with two values of interest rate: $r = 0.02 < \lambda$ and $r = 0.06 > \lambda$. We plot the robust ruin probability, the

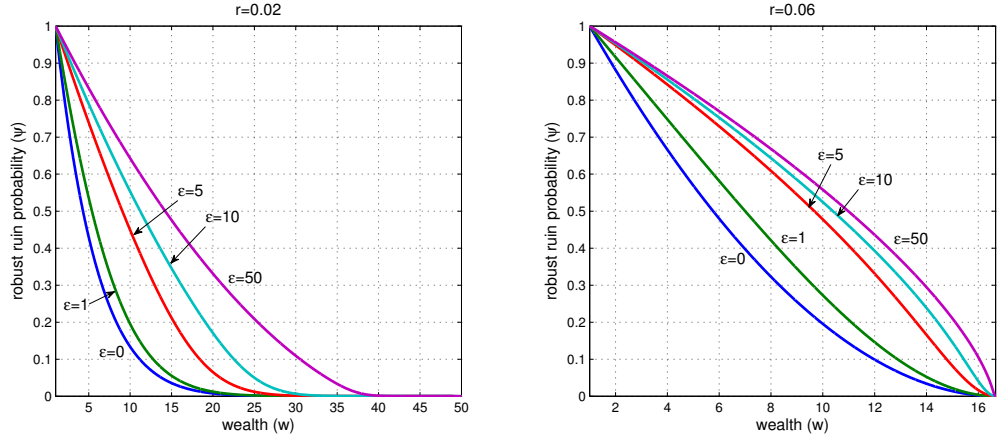


Figure 3.2: Robust ruin probabilities.

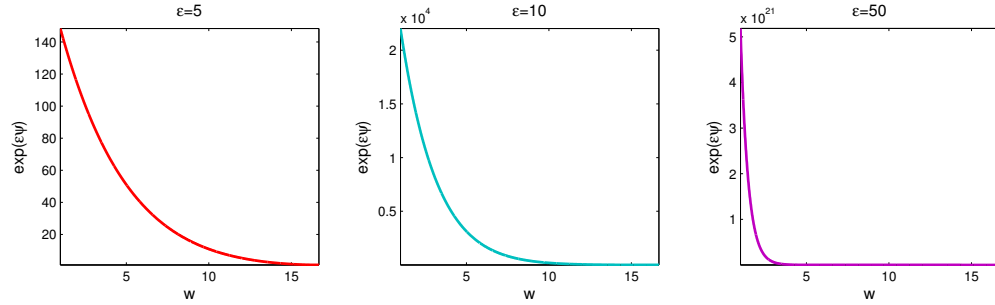


Figure 3.3: Cole-Hopf transform of three non-convex curves in Figure 3.2.

optimal investment and the optimal distorted Sharpe ratio as functions of wealth under different levels of ambiguity aversion.

From Figure 3.2, we see that the robust value function is increasing in ε . When the interest rate is smaller than the hazard rate, all value functions are strictly convex. When the interest rate is larger than the hazard rate, concavity depends on the level of ambiguity aversion: the value function is convex when ε is small, and changes from concave to convex when ε is large. The larger the ε , the closer the inflection point is to the safe level. With this set of parameters, a sufficient condition for ψ to be convex, as implied by Proposition 3.7.1, is $0 \leq \varepsilon \leq 0.4765$. A sufficient condition for ψ to change concavity is $\varepsilon > 1.7778$. $\varepsilon = 5, 10, 50$ all satisfy this condition and

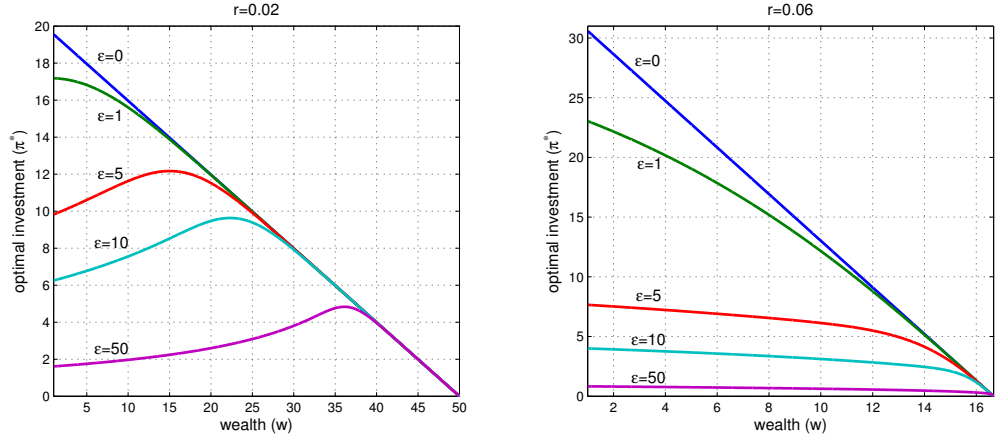


Figure 3.4: Optimal investments.

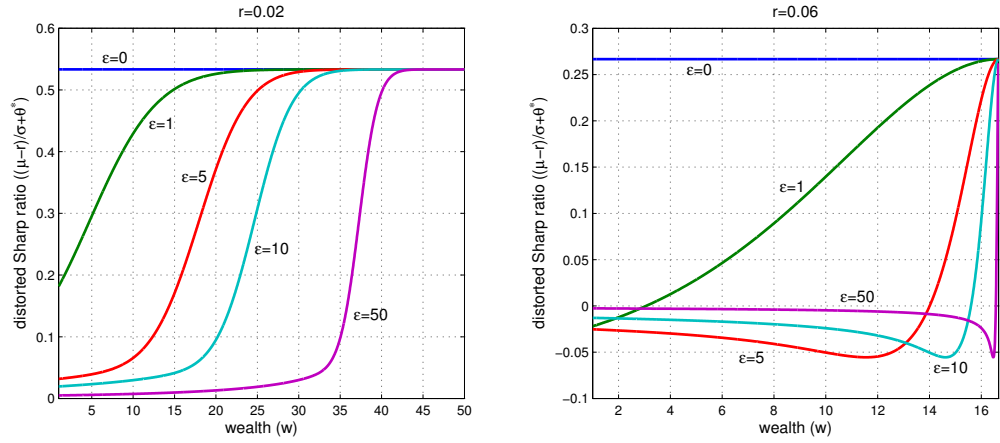


Figure 3.5: Distorted Sharpe ratios.

exhibit concavity change. By Remark 3.7.2, the inflection points of ψ corresponds to the points where the optimal distorted Sharpe ratio is zero. Despite that ψ may be concave, its Cole-Hopf transform $e^{\varepsilon\psi}$ is always convex, as demonstrated by Figure 3.3.

Figure 3.4 shows the optimal investment level is decreasing in ε , which agrees with Proposition 3.6.3. This means the more ambiguity-averse the agent is, the less she is willing to invest in the risky asset. Different from the non-robust case, the

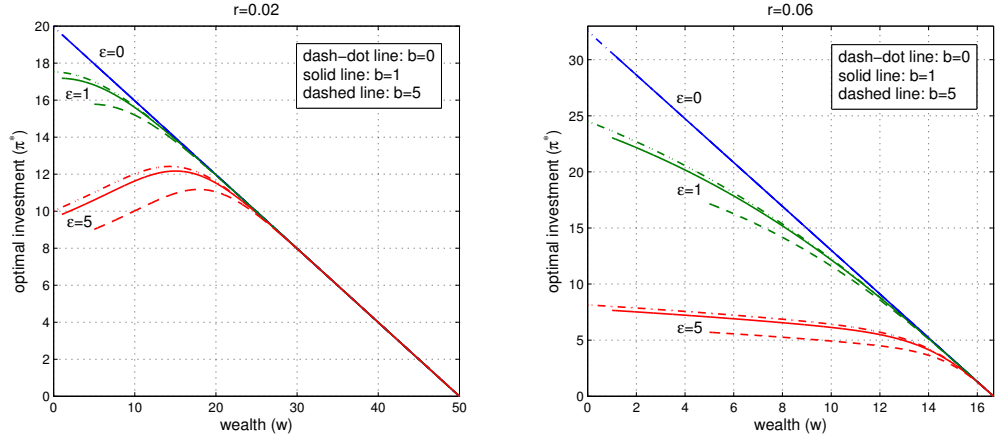


Figure 3.6: Optimal investments with different ruin levels.

optimal investment, although goes to zero as wealth approaches the safe level, is not necessarily a decreasing function of wealth. When interest rate is large (compared with hazard rate), π^* is decreasing and also concave in w . But when interest rate is small, there is an interior point where π^* achieves maximum. Moreover, as ε increases, the maximum point moves to the right. In any case, adding ambiguity aversion reduces the amount of borrowing when the wealth of the investor is small, making the model more realistic than the non-robust model (without borrowing constraint). Another interesting observation is that the optimal π^* of all levels of ambiguity aversion share the same tangent line at the safe level with their non-robust counterparts, confirming equation (3.30) of Proposition 3.6.4.

Figure 3.5 shows that when the interest rate is small (compared with hazard rate), the optimally distorted Sharpe ratio $\frac{\mu-r}{\sigma} + \theta^*$ is strictly positive, decreasing in ε and increasing in wealth. But when interest rate is large, it can be negative, and both monotonicities are lost. In both cases, the pictures suggest that the optimally distorted Sharpe ratio is converging to zero pointwise as $\varepsilon \rightarrow \infty$. Moreover, observe from Figure 3.4 that the optimal investment converges to zero pointwise as $\varepsilon \rightarrow \infty$,

which is the investment behavior corresponding to ψ_∞ . This suggests that for ε very large, the stock is losing its attractiveness as it becomes less favorable compared to the money market account.

In the non-robust case, the optimal investment strategy is independent of the ruin level b in the sense that if $b_1 < b_2$, then $\pi^*(w; b_1)$ coincides with $\pi^*(w; b_2)$ on $[b_2, w_s]$. This holds not only for constant consumption rate, but also for any Lipschitz continuous consumption rate (see [13, Corollary 2.3]). However, when ambiguity aversion is present, the ruin level has a global impact on the investment decision unless hazard rate is zero. When $\varepsilon \neq 0$, Figure 3.6 suggests π^* is decreasing in b . In other words, the individual will invest less if she is more likely to feel ruined.

3.9.1 Should the individual care about robustness?

Figure 3.2 shows robustness has a considerable impact on the minimum probability of ruin. However, this is not really informative as far as investment behavior is concerned. A more important question is: how does the optimal non-robust investment strategy π_0 perform in the robust market? In other words, will the individual bear significantly more risk if she makes investment decisions as if there were no model uncertainty? The answer to this question is partially affirmative; the individual should care about robustness for non-small ε . In our numerical example, ignoring robustness increases the ruin probability by more than 10% for ε larger than 10. On the other hand, for small ε , π_0 turns out to be a good enough investment strategy. For $\varepsilon = 1$, the difference between the ruin probability yielded by π_0 and the optimal ruin probability $\psi(\cdot; 1)$ is on a scale of 0.1% which may be negligible for an individual, although the difference between ψ_0 and $\psi(\cdot; 1)$ can be as large as 10%. Table 3.1 illustrates the performance of π_0 under various levels of ambiguity aversion in

our numerical example.

Table 3.1: Maximum deviation from the minimum robust ruin probability if the individual uses the non-robust strategy π_0 .

	$\varepsilon = 1$	$\varepsilon = 2$	$\varepsilon = 3$	$\varepsilon = 4$	$\varepsilon = 5$	$\varepsilon = 10$	$\varepsilon = 20$
$r = 0.02$	0.001	0.005	0.013	0.025	0.038	0.105	0.201
$r = 0.06$	0.002	0.013	0.033	0.059	0.087	0.198	0.324

CHAPTER IV

Stochastic Perron's Method for the Lifetime Ruin Problem Under Transaction Costs

4.1 Introduction

Stochastic Perron's method is introduced in [9], [11] and [10] as a way to show the value function of a stochastic control problem is the unique viscosity solution of the associated Hamilton-Jacobi-Bellman (HJB) equation, without having to first go through the proof of the dynamic programming principle (DPP) which is usually very long and complicated, and often incomplete. It is a direct verification approach in that it first constructs a solution to the HJB equation, and then verifies such a solution is the value function. But unlike the classical verification, it does not require regularity; uniqueness acts as a substitute for verification. The basic idea is to define, for each specific problem, a suitable family of stochastic supersolutions \mathcal{V}^+ (resp. stochastic subsolutions \mathcal{V}^-) which is stable under minimum (resp. maximum), and whose members bound the value function from above (resp. below). So the value function is enveloped from above by $v_+ = \inf_{v \in \mathcal{V}^+} v$ and from below by $v_- = \sup_{v \in \mathcal{V}^-} v$. The key step is to show v_+ is a viscosity subsolution and v_- is a viscosity supersolution by a Perron-type argument. A comparison principle then closes the gap.

Stochastic Perron's method has been applied to linear problems [9], Dynkin games [11], HJB equations for regular control problems [10], (regular) exit time problems [68], zero-sum differential games [74], and after our paper came out, control problems with state constraints [70], robust switching problems [6], stochastic target games [8] and regular finite fuel problems [69]. This chapter adapts the method to another type of problems: singular control problems. In particular, we focus on the specific problem of how individuals should invest their wealth in a risky financial market to minimize the probability of lifetime ruin, when buying and selling of the risky asset incur proportional transaction costs. This problem can also be treated as an exit time problem, but with singular controls. In the frictionless case, the probability of lifetime ruin problem was analyzed by Young [82], and later studied in more complicated settings such as borrowing constraints, stochastic consumption and drift uncertainty (see Chapter III for a detailed introduction). The main goal in this chapter is to exemplify how stochastic Perron's method can be applied to singular control problems, which has not been covered in the literature. It also serves as the first step towards a rigorous analysis of the probability of lifetime ruin problem under transaction costs. The techniques in this paper can be applied in a similar way to other optimal investment problems under transaction costs, as long as there is a comparison principle. For consumption-investment problems, uniqueness is proved in [53] under certain conditions (also see [79, Theorem 1] and Section 4.3 of [55]).

The main idea of the proof is in line with [10] and [68], but there are some nontrivial modifications. Similar to [32] and [73], our HJB equation takes the form of a variational inequality with three components, one for each of the three different regions: no-transaction, sell, and buy. This makes the proof of the interior viscosity subsolution property of the upper stochastic envelope v_+ more demanding: we have

to argue by contradiction in three cases separately. Variational inequalities also appear in [11] and the authors are able to rule out some of the cases by assuming the existence of a stochastic supersolution (resp. subsolution) less than or equal to the upper obstacle (resp. greater than or equal to the lower obstacle). But the same idea does not work for gradient constraints. Another challenge posed by the singular control is that the state process can jump outside the small neighborhood in which local estimates obtained from the viscosity solution property are valid. This issue arises in the proof of the interior viscosity supersolution property of the lower stochastic envelope v_- , and we overcome it by splitting the jump into two steps: first to an intermediate point on the boundary of the neighborhood and then to its original destination.

In proving the viscosity semi-solution property of v_{\pm} , boundary property is usually harder to show than interior property. In fact, most of the work in [68] is devoted to proving the boundary viscosity semi-solution property of v_{\pm} . In our case, we avoid this hassle by constructing explicitly a stochastic supersolution and a stochastic subsolution both of which satisfy the boundary condition. The boundary viscosity semi-solution property then becomes a trivial consequence of the definition of v_{\pm} . This is very similar to classical Perron's method in which one has to first come up with a pair of viscosity semi-solutions satisfying the boundary condition (see Theorem 4.1 and Example 4.6 of [29]). However, we point out that the construction of such stochastic semi-solutions depends on the specific problem at hand and may not always be possible.

Previous works on stochastic Perron's method focus on methodology and take comparison principle (which is crucial for stochastic Perron's method to work) as an assumption. Here we provide, in addition to stochastic Perron's method, a complete

proof of the comparison principle for our specific singular control problem. The proof relies on the existence of a strict classical subsolution satisfying certain growth condition, an idea we borrowed from [53].

The rest of this chapter is organized as follows. In Section 4.2, we set up the problem, derive the HJB equation and some bounds on the value function, and state the main theorem. In Section 4.3, we introduce the notion of stochastic supersolution and show the infimum of stochastic supersolutions is a viscosity subsolution. In Section 4.4, we introduce the notion of stochastic subsolution and show the supremum of stochastic subsolutions is a viscosity supersolution. Finally, in Section 4.5 we prove a comparison principle and finish the proof of the main theorem.

4.2 Problem formulation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting a Brownian motion $W = (W_t)_{t \geq 0}$ and an independent Poisson process $N = (N_t)_{t \geq 0}$ with rate β . Let τ_d be the first time that the Poisson process jumps, modeling the death time of the individual. τ_d is exponentially distributed with rate β , known as the hazard rate in this context. Denote by $\mathbb{F} := \{\mathcal{F}_t\}_{t \geq 0}$ the completion of the natural filtration of the Brownian motion and $\mathbb{G} := \{\mathcal{G}_t\}_{t \geq 0}$ the completion of the filtration generated by W and the process $1_{\{t \geq \tau_d\}}$. Assume both \mathcal{F} and \mathcal{G} have been made right continuous; that is, they satisfy the usual condition.

The financial market consists of a risk-free money market with interest rate $r > 0$ and a risky asset (a stock) whose price P_t follows a geometric Brownian motion with drift $\alpha > r$ and volatility $\sigma > 0$. Transferring assets between the money market and the stock market incur proportional transaction costs specified by two parameters $\lambda, \mu \in (0, 1)$. One can think of the stock as having ask price $P_t/(1 - \lambda)$ and bid

price $(1 - \mu)P_t$. Same as [73], we describe the investment policy of the individual by a pair (B, S) of right-continuous with left limits (RCLL), non-negative, non-decreasing and \mathbb{G} -adapted processes, where B records the cumulative amount of money withdrawn from the money market for the purpose of buying stock, and S records the cumulative sales of stock for the purpose of investment in the money market. We set $(B_{0-}, S_{0-}) = \mathbf{0}$, i.e. there is no investment history at time zero. Due to transaction costs, it is never optimal to buy and sell at the same time. So we limit ourselves to strategies (B, S) such that for all t , $\Delta B_t := B_t - B_{t-}$ and $\Delta S_t := S_t - S_{t-}$ are not both strictly positive. Denote by \mathcal{A}_0 the set of all such pairs (B, S) . Apart from investment, the individual also consumes at a constant rate $c > 0$.

Denote by X_t and Y_t the total dollar amount invested in the money market and the stock at time t , respectively. Let $L(x, y) := x + (1 - \mu)y^+ - \frac{1}{1-\lambda}y^-$ be the liquidation function. For each $a \in \mathbb{R}$, define

$$\mathcal{S}_a := \{(x, y) \in \mathbb{R}^2 : L(x, y) > a\} = \{(x, y) \in \mathbb{R}^2 : x + \frac{y}{1-\lambda} > a, x + (1 - \mu)y > a\}.$$

Given initial endowment (x, y) and a pair of control $(B, S) \in \mathcal{A}_0$, the pre-death investment position of the individual evolve according to the stochastic differential equations (SDE)

$$(4.1) \quad dX_t = (rX_t - c)dt - dB_t + (1 - \mu)dS_t, \quad X_{0-} = x,$$

$$(4.2) \quad dY_t = \alpha Y_t dt + \sigma Y_t dW_t + (1 - \lambda)dB_t - dS_t, \quad Y_{0-} = y.$$

Here we allow an immediate transaction at time zero so that (X_0, Y_0) may differ from (x, y) . Denote the solution by $(X^{x,y,B,S}, Y^{x,y,B,S})$. Let

$$\tau_b^{x,y,B,S} := \inf\{t \geq 0 : (X^{x,y,B,S}, Y^{x,y,B,S}) \notin \mathcal{S}_b\}$$

be the ruin time. The individual aims at minimizing the probability that ruin hap-

pens before death. The value function of this control problem is defined as

$$(4.3) \quad \psi(x, y) := \inf_{(B, S) \in \mathcal{A}_0} \mathbb{P}(\tau_b^{x, y, B, S} < \tau_d).$$

Clearly, ψ is $[0, 1]$ -valued, and $\psi(x, y) = 1$ if $(x, y) \notin \mathcal{S}_b$. Same as in the frictionless case, when $L(x, y) \geq c/r$, the individual can sustain her consumption by immediately putting all her money in the money market and consuming the interest. We shall assume $b < c/r$, otherwise the problem is trivial.¹ We have $\psi(x, y) = 0$ for $(x, y) \in \overline{\mathcal{S}}_{c/r}$. In other words, $\overline{\mathcal{S}}_{c/r}$ is a “safe region”. The (open) state space for this control problem is $\mathcal{S} := \mathcal{S}_b \setminus \overline{\mathcal{S}}_{c/r}$, and the boundary consists of two parts: the ruin level $\partial\mathcal{S}_b$ and the safe level $\partial\mathcal{S}_{c/r}$.

For $\varphi \in C^2(\mathcal{S})$, define

$$\mathcal{L}\varphi := \beta\varphi - (rx - c)\varphi_x - \alpha y\varphi_y - \frac{1}{2}\sigma^2 y^2 \varphi_{yy}.$$

The HJB equation for the frictional lifetime ruin problem is

$$(4.4) \quad \max \{ \mathcal{L}u, -(1 - \mu)u_x + u_y, u_x - (1 - \lambda)u_y \} = 0, \quad (x, y) \in \mathcal{S},$$

with boundary conditions

$$(4.5) \quad u(x, y) = 1 \text{ if } (x, y) \in \partial\mathcal{S}_b, \quad u(x, y) = 0 \text{ if } (x, y) \in \partial\mathcal{S}_{c/r}.$$

4.2.1 Upper and lower bounds on the value function

Let

$$(4.6) \quad \overline{\psi}(x, y) := \left(\frac{c - rL(x, y)}{c - rb} \right)^{\frac{\beta}{r}}, \quad (x, y) \in \overline{\mathcal{S}}.$$

$\overline{\psi}$ is the probability of ruin if the agent immediately liquidate her stock position and makes no further transaction throughout her lifetime. It is an upper bound for the

¹If $b \geq c/r$, then $\psi(x, y)$ is either 0 or 1, depending on whether (x, y) belongs to \mathcal{S}_b or not.

value function since such a strategy may not be optimal. It is easy to see that $\bar{\psi}$ satisfies the boundary conditions (4.5).

For $k \in [1 - \mu, \frac{1}{1-\lambda}]$, let

$$(4.7) \quad \psi_k(x, y) := \begin{cases} \left(\frac{c - r(x + ky)}{c - rb} \right)^d, & b \leq x + ky \leq c/r, \\ 0, & x + ky > c/r. \end{cases}$$

where

$$(4.8) \quad d = \frac{1}{2r} \left[(r + \beta + R) + \sqrt{(r + \beta + R)^2 - 4r\beta} \right] > 1, \quad R = \frac{1}{2} \left(\frac{\alpha - r}{\sigma} \right)^2.$$

That is, $\psi_k(x, y)$ is the minimum frictionless probability of ruin when the initial wealth is $x + ky$ (the frictionless ruin probability is derived in [82]). ψ_k bounds the frictional value function from below because each k corresponds to a stock price inside the bid-ask spread, and trading at a more favorable frictionless price obviously leads to smaller ruin probability. For a rigorous proof, one can refer to Remark 4.4.3 and Lemma 4.4.6. Since the value function ψ is bounded from below by ψ_k for each k , it is bounded from below by their supremum:

$$(4.9) \quad \underline{\psi}(x, y) := \sup_{k \in [1 - \mu, \frac{1}{1-\lambda}]} \psi_k(x, y) = \psi_{1-\mu}(x, y) \vee \psi_{\frac{1}{1-\lambda}}(x, y) = \left(\frac{c - rL(x, y)}{c - rb} \right)^d.$$

Since ψ_k is continuous in k , the above supremum remains unchanged if we replace $[1 - \mu, \frac{1}{1-\lambda}]$ by $(1 - \mu, \frac{1}{1-\lambda}) \cap \mathbb{Q}$. Clearly, $\underline{\psi}$ satisfies the boundary conditions (4.5).

The following lemma summarizes the results.

Lemma 4.2.1. *For $(x, y) \in \bar{\mathcal{S}}$,*

$$\left(\frac{c - rL(x, y)}{c - rb} \right)^d \leq \psi(x, y) \leq \left(\frac{c - rL(x, y)}{c - rb} \right)^{\frac{\beta}{r}},$$

where d is defined in (4.8).

Remark 4.2.2. It can be shown that $\bar{\psi}$ is a viscosity supersolution and $\underline{\psi}$ is a viscosity subsolution of (4.4). With a comparison principle which we will prove in Section 4.5, one can use (classical) Perron's method (see Chapter III, [46] or [29]) to get the existence of a viscosity solution to (4.4), (4.5). But such a solution cannot be compared with the value function unless one can prove regularity which is necessary for the classical verification theorem. Instead, we will use stochastic Perron's method which amounts to verification without smoothness.

4.2.2 Random initial condition and admissible controls

For convenience in later discussion, we introduce a ‘‘coffin state’’ Δ . Let $\bar{\mathcal{S}} \cup \Delta$ be the one point compactification of $\bar{\mathcal{S}}$. Throughout this paper, all closures are taken in \mathbb{R}^2 . For any \mathbb{R}^2 -valued vector z , we use the convention that $\Delta + z = \Delta$. Set $(X_t, Y_t) := \Delta$ for all $t \geq \tau_d$. For any function u defined on $\bar{\mathcal{S}}$, define its extension to $\bar{\mathcal{S}} \cup \{\Delta\}$ by assigning $u(\Delta) = 0$.

A pair (τ, ξ) is called a *random initial condition* for (4.1), (4.2) if τ is a \mathbb{G} -stopping time taking values in $[0, \tau_d]$, $\xi = (\xi^0, \xi^1)$ is a \mathcal{G}_τ -measurable random vector taking values in $\bar{\mathcal{S}} \cup \{\Delta\}$, and $\xi = \Delta$ if and only if $\tau = \tau_d$. Denote by $(X^{\tau, \xi, B, S}, Y^{\tau, \xi, B, S})$ the solution of (4.1) and (4.2) with random initial condition (τ, ξ) in the sense that $(X_{\tau-}, Y_{\tau-}) = \xi$. The *exit time* of $(X^{\tau, \xi, B, S}, Y^{\tau, \xi, B, S})$ from \mathcal{S} is defined by

$$\sigma^{\tau, \xi, B, S} := \inf\{t \geq \tau : (X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S}) \notin \mathcal{S}\}.$$

Note that $\sigma^{\tau, \xi, B, S} \leq \tau_d < \infty$ since $(X_{\tau_d}^{\tau, \xi, B, S}, Y_{\tau_d}^{\tau, \xi, B, S}) = \Delta \notin \mathcal{S}$.

We also restrict ourselves to a subset of controls. Observe that when buying stocks, we move northwest along the vector $(-1, 1 - \lambda)$; when selling stocks, we move southeast along the vector $(1 - \mu, -1)$. It is not hard to see by picture that starting in \mathcal{S} , one can never jump to $\mathcal{S}_{c/r}$ by a transaction. On the other hand, it is never

optimal to jump across $\partial\mathcal{S}_b$ from \mathcal{S} because such a jump immediately leads to ruin. If we are on $\partial\mathcal{S}_{c/r}$ (resp. $\partial\mathcal{S}_b$), jumping to its right is impossible and jumping to its left is not optimal (resp. does not prevent ruin from happening). Therefore, we may focus on those controls under which the controlled process exits \mathcal{S} via its boundary or the coffin state. The formal definition of admissibility is given below.

Definition 4.2.3. Let (τ, ξ) be a random initial condition. A control pair $(B, S) \in \mathcal{A}_0$ is called (τ, ξ) -admissible if

$$(X_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}, Y_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}) \in \partial\mathcal{S} \cup \{\Delta\}.$$

Denote the set of (τ, ξ) -admissible controls by $\mathcal{A}(\tau, \xi)$.

We have $(B, S) \equiv \mathbf{0} \in \mathcal{A}(\tau, \xi)$ for any random initial condition (τ, ξ) . When $\tau = 0$ and $\xi = (x, y)$, we shall omit the τ -dependence in the superscripts of the controlled process and relevant stopping times, and write $\mathcal{A}(\tau, \xi) = \mathcal{A}(x, y)$. As we have argued, working with admissible controls does not change the optimal probability, i.e.

$$\psi(x, y) = \inf_{(B, S) \in \mathcal{A}(x, y)} \mathbb{P}(\tau_b^{x, y, B, S} < \tau_d).$$

The following constructions of admissible controls will be used a few times in Section 4.3. We list them here for future reference.

Lemma 4.2.4.

(i) If (B^i, S^i) , $i = 1, 2$ are (τ, ξ) -admissible and A is any \mathcal{G}_τ -measurable set, then

$$(B_t, S_t) := 1_{\{t \geq \tau\}} [(B_t^1 - B_{\tau-}^1, S_t^1 - S_{\tau-}^1) 1_A + (B_t^2 - B_{\tau-}^2, S_t^2 - S_{\tau-}^2) 1_{A^c}]$$

is also (τ, ξ) -admissible.

(ii) Let (B^1, S^1) be a (τ, ξ) -admissible control, $\tau_1 \in [\tau, \sigma^{\tau, \xi, B^1, S^1}]$ be a \mathbb{G} -stopping time, and $\xi_1 := (X_{\tau_1}^{\tau, \xi, B^1, S^1}, Y_{\tau_1}^{\tau, \xi, B^1, S^1})$. Then (τ_1, ξ_1) is a random initial condition. Furthermore, let (B^2, S^2) be a (τ_1, ξ_1) -admissible control. Then

$$(B_t, S_t) := 1_{\{t < \tau_1\}}(B_t^1, S_t^1) + 1_{\{t \geq \tau_1\}}(B_t^2 - B_{\tau_1-}^2 + B_{\tau_1}^1, S_t^2 - S_{\tau_1-}^2 + S_{\tau_1}^1)$$

is a (τ, ξ) -admissible control.

Proof. (i) (B, S) is \mathbb{G} -adapted by the definition of stopping time and stopping time filtration, and the \mathbb{G} -adaptedness of (B^i, S^i) , $i = 1, 2$. It is nonnegative because (B^i, S^i) , $i = 1, 2$ are non-decreasing. Monotonicity, RCLL property and that ΔB and ΔS are not both strictly positive also follow from the assumption that $(B^i, S^i) \in \mathcal{A}_0$, $i = 1, 2$. So $(B, S) \in \mathcal{A}_0$. By pathwise uniqueness of the solution to (4.1), (4.2), we have

$$(X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S}) = 1_A(X_t^{\tau, \xi, B^1, S^1}, Y_t^{\tau, \xi, B^1, S^1}) + 1_{A^c}(X_t^{\tau, \xi, B^2, S^2}, Y_t^{\tau, \xi, B^2, S^2}), \quad t \geq \tau.$$

It follows that

$$\sigma^{\tau, \xi, B, S} = 1_A \sigma^{\tau, \xi, B^1, S^1} + 1_{A^c} \sigma^{\tau, \xi, B^2, S^2},$$

and thus

$$(X_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}, Y_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}) = 1_A(X_{\sigma^{\tau, \xi, B^1, S^1}}^{\tau, \xi, B^1, S^1}, Y_{\sigma^{\tau, \xi, B^1, S^1}}^{\tau, \xi, B^1, S^1}) + 1_{A^c}(X_{\sigma^{\tau, \xi, B^2, S^2}}^{\tau, \xi, B^2, S^2}, Y_{\sigma^{\tau, \xi, B^2, S^2}}^{\tau, \xi, B^2, S^2}) \in \partial \mathcal{S} \cup \{\Delta\}$$

by the (τ, ξ) -admissibility of (B^i, S^i) , $i = 1, 2$.

(ii) Clearly, τ_1 is a \mathbb{G} -stopping time taking values in $[\tau, \tau_d]$ and ξ_1 is \mathcal{G}_{τ_1} -measurable. Since $\tau_1 \leq \sigma^{\tau, \xi, B^1, S^1}$, the (τ, ξ) -admissibility of (B^1, S^1) implies $\xi_1 \in \bar{\mathcal{S}} \cup \{\Delta\}$. Moreover, $\xi_1 = \Delta$ if and only if $\tau_1 = \tau_d$. So (τ_1, ξ_1) is a valid random initial condition. It is routine to check $(B, S) \in \mathcal{A}_0$. To show $(B, S) \in \mathcal{A}(\tau, \xi)$, observe that

$$(X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S}) = \begin{cases} (X_t^{\tau, \xi, B^1, S^1}, Y_t^{\tau, \xi, B^1, S^1}), & \tau \leq t < \tau_1, \\ (X_t^{\tau_1, \xi_1, B^2, S^2}, Y_t^{\tau_1, \xi_1, B^2, S^2}), & t \geq \tau_1. \end{cases}$$

This, together with $\tau_1 \leq \sigma^{\tau, \xi, B^1, S^1}$, imply $\sigma^{\tau, \xi, B, S} = \sigma^{\tau_1, \xi_1, B^2, S^2} \geq \tau_1$. Since $(B^2, S^2) \in \mathcal{A}(\tau_1, \xi_1)$, we have

$$(X_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}, Y_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}) = (X_{\sigma^{\tau_1, \xi_1, B^2, S^2}}^{\tau_1, \xi_1, B^2, S^2}, Y_{\sigma^{\tau_1, \xi_1, B^2, S^2}}^{\tau_1, \xi_1, B^2, S^2}) \in \partial\mathcal{S} \cup \{\Delta\}.$$

□

4.2.3 Main results

Theorem 4.2.5. *The value function ψ is the unique (continuous) viscosity solution to the HJB equation (4.4) satisfying the boundary condition (4.5).*

The proof of Theorem 4.2.5 is deferred to the end of Section 4.5.

4.3 Stochastic supersolution

Definition 4.3.1. A bounded u.s.c. function v on $\bar{\mathcal{S}}$ is called a *stochastic supersolution* of (4.4), (4.5) if

(SP1) $v \geq 1$ on $\partial\mathcal{S}_b$, $v \geq 0$ on $\partial\mathcal{S}_{c/\tau}$;

(SP2) for any random initial condition (τ, ξ) , there exists $(B, S) \in \mathcal{A}(\tau, \xi)$ such that

$$\mathbb{E}[v(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \leq v(\xi)$$

for all \mathbb{G} -stopping time $\rho \in [\tau, \sigma^{\tau, \xi, B, S}]$, where v is understood to be its extension to $\bar{\mathcal{S}} \cup \{\Delta\}$.

Denote the set of stochastic supersolutions by \mathcal{V}^+ .

Remark 4.3.2. $\mathcal{V}^+ \neq \emptyset$ since the constant 1 $\in \mathcal{V}^+$. There is a more useful stochastic supersolution: the upper bound function $\bar{\psi}$ defined in (4.6), which satisfies (SP1) with equality. (See Lemma 4.3.4.) The existence of such a stochastic supersolution automatically guarantees the the upper stochastic envelope (which will be introduced shortly) satisfies the boundary condition (4.5).

Remark 4.3.3. Any stochastic supersolution v dominates the value function ψ on $\bar{\mathcal{S}}$. To see this, first note that $v \geq \psi$ on $\partial\mathcal{S}$ by (SP1). Then for any $(x, y) \in \mathcal{S}$, take $\tau = 0$ and $\xi = (x, y)$. Let $(B, S) \in \mathcal{A}(x, y)$ be given by (SP2) for v . Let $\rho = \sigma^{x,y,B,S}$. To simplify notation, we write τ_b for $\tau_b^{x,y,B,S}$ and τ_s for $\tau_s^{x,y,B,S} := \inf\{t \geq 0 : (X_t^{x,y,B,S}, Y_t^{x,y,B,S}) \in \bar{\mathcal{S}}_{c/r}\}$. We have

$$v(x, y) \geq \mathbb{E} [v(X_\rho^{x,y,B,S}, Y_\rho^{x,y,B,S})] \geq \mathbb{E} \left[1_{\{(X_\rho^{x,y,B,S}, Y_\rho^{x,y,B,S}) \in \partial\mathcal{S}_b\}} \right] = \mathbb{P}(\tau_b < \tau_d \wedge \tau_s).$$

where the first inequality holds by (SP2) and the second inequality holds by (SP1).

Now, let

$$(B'_t, S'_t) = (B_t, S_t)1_{\{t < \tau_s\}} + ((X_{\tau_s}^{x,y,B,S} - c/r)^+ + B_{\tau_s}, (Y_{\tau_s}^{x,y,B,S})^+ + S_{\tau_s})1_{\{t \geq \tau_s\}}.$$

That is, (B', S') follows (B, S) before hitting the safe region, and at the moment when the safe region is hit (by diffusion), immediately liquidate all stock position and do no more transaction afterwards. This ensures that once the safe region is reached, death will definitely happen before ruin. It is easy to check $(B', S') \in \mathcal{A}_0$ and $\mathbb{P}(\tau_b < \tau_d \wedge \tau_s) = \mathbb{P}(\tau_b^{x,y,B',S'} < \tau_d)$. We therefore have

$$v(x, y) \geq \mathbb{P}(\tau_b^{x,y,B',S'} < \tau_d) \geq \psi(x, y).$$

Lemma 4.3.4. $\bar{\psi} \in \mathcal{V}^+$.

Proof. We only show (SP2). Let (τ, ξ) be any random initial condition. Define

$$(B_t, S_t) := 1_{\{t \geq \tau\}} \left(\frac{(\xi^1)^-}{1 - \lambda}, (\xi^1)^+ \right).$$

Intuitively, what (B, S) does is to immediately liquidate the stock position at time τ and do no more transaction afterwards. It can be checked that $(X_{\sigma^{\tau,\xi,B,S}}^{\tau,\xi,B,S}, Y_{\sigma^{\tau,\xi,B,S}}^{\tau,\xi,B,S}) \in \{(b, 0), (c/r, 0), \Delta\}$, thus $(B, S) \in \mathcal{A}(\tau, \xi)$. We have

$$(X_\tau^{\tau,\xi,B,S}, Y_\tau^{\tau,\xi,B,S}) = 1_{\{\tau < \tau_d\}}(L(\xi), 0) + 1_{\{\tau = \tau_d\}}\Delta,$$

and $Y_t^{\tau,\xi,B,S} = 0$ for all $t \in [\tau, \tau_d)$. Let $\rho \in [\tau, \sigma^{\tau,\xi,B,S}]$ be any \mathbb{G} -stopping time. Let $f(x) := \bar{\psi}(x, 0) \in C[b, c/r] \cap C^2[b, c/r]$. With slight abuse of notation, we also write $X_t^{\tau,\xi,B,S} = \Delta$ when $t = \tau_d$, and set $f(\Delta) = 0$. Apply Itô's formula to $f(X^{\tau,\xi,B,S})$, we get

$$\begin{aligned}
& f(X_\rho^{\tau,\xi,B,S}) - f(X_\tau^{\tau,\xi,B,S}) \\
&= \int_\tau^\rho f'(X_t^{\tau,\xi,B,S})(rX_t^{\tau,\xi,B,S} - c)dt + \int_\tau^\rho \left(f(\Delta) - f(X_{t-}^{\tau,\xi,B,S}) \right) dN_t \\
&= \int_\tau^\rho [f'(x)(rx - c) - \beta f(x)] \Big|_{x=X_t^{\tau,\xi,B,S}} dt + \int_\tau^\rho -f(X_{t-}^{\tau,\xi,B,S})d(N_t - \beta t) \\
&= \int_\tau^\rho -f(X_{t-}^{\tau,\xi,B,S})d(N_t - \beta t),
\end{aligned}$$

where we used the explicit formula of f to kill the drift. Taking conditional expectation yields

$$\mathbb{E}[f(X_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] = f(X_\tau^{\tau,\xi,B,S}).$$

It follows that

$$\begin{aligned}
\mathbb{E}[\bar{\psi}(X_\rho^{\tau,\xi,B,S}, Y_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] &= \mathbb{E}[1_{\{\rho < \tau_d\}}\bar{\psi}(X_\rho^{\tau,\xi,B,S}, 0)|\mathcal{G}_\tau] = \mathbb{E}[1_{\{\rho < \tau_d\}}f(X_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] \\
&= \mathbb{E}[f(X_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] = f(X_\tau^{\tau,\xi,B,S}) = 1_{\{\tau < \tau_d\}}f(L(\xi)) \\
&= 1_{\{\tau < \tau_d\}}\bar{\psi}(L(\xi), 0) = 1_{\{\tau < \tau_d\}}\bar{\psi}(\xi) = \bar{\psi}(\xi).
\end{aligned}$$

In the second last equality, we used $\bar{\psi}(x, y) = \bar{\psi}(L(x, y), 0)$ for $(x, y) \in \bar{\mathcal{S}}$. \square

Lemma 4.3.5. *Let $v_1, v_2 \in \mathcal{V}^+$. Then $v_1 \wedge v_2 \in \mathcal{V}^+$.*

Proof. The minimum of bounded u.s.c. functions is still bounded and u.s.c.. (SP1) is clearly satisfied. For (SP2), let $(B^i, S^i) \in \mathcal{A}(\tau, \xi), i = 1, 2$ be the admissible control corresponding to v_i and the random initial condition (τ, ξ) . Put $A := \{v_1(\xi) \leq v_2(\xi)\} \in \mathcal{G}_\tau$. The control

$$(B_t, S_t) := 1_{\{t \geq \tau\}} [(B_t^1 - B_{\tau-}^1, S_t^1 - S_{\tau-}^1) 1_A + (B_t^2 - B_{\tau-}^2, S_t^2 - S_{\tau-}^2) 1_{A^c}]$$

serves the purpose. (τ, ξ) -admissible follows from Lemma 4.2.4.i, and the remaining proof is very similar to that of [68, Lemma 1] except that the process Z is replaced by $v(X, Y)$ and the direction of inequalities are reversed. So we omit the details. \square

Proposition 4.3.6. *The upper stochastic envelope*

$$v_+(x, y) := \inf_{v \in \mathcal{V}^+} v(x, y)$$

is a viscosity subsolution of (4.4) satisfying $v_+ \leq 1$ on $\partial\mathcal{S}_b$ and $v_+ \leq 0$ on $\partial\mathcal{S}_{c/r}$.

Proof. The boundary inequalities are satisfied because $v_+ \leq \bar{\psi}$ by Lemma 4.3.4.² To show interior viscosity subsolution property, let $(x_0, y_0) \in \mathcal{S}$ and $\varphi \in C^2(\mathcal{S})$ be a test function such that $v_+ - \varphi$ attains a strict local maximum of zero at (x_0, y_0) . We need to show

$$\max\{\mathcal{L}\varphi, -(1 - \mu)\varphi_x + \varphi_y, \varphi_x - (1 - \lambda)\varphi_y\}(x_0, y_0) \leq 0.$$

Assume on the contrary that

$$\max\{\mathcal{L}\varphi, -(1 - \mu)\varphi_x + \varphi_y, \varphi_x - (1 - \lambda)\varphi_y\}(x_0, y_0) > 0.$$

There are three cases to consider: (i) $\mathcal{L}\varphi(x_0, y_0) > 0$, (ii) $-(1 - \mu)\varphi_x(x_0, y_0) + \varphi_y(x_0, y_0) > 0$, and (iii) $\varphi_x(x_0, y_0) - (1 - \lambda)\varphi_y(x_0, y_0) > 0$. We will show that each case leads to a contradiction.

Case (i). $\mathcal{L}\varphi(x_0, y_0) > 0$. We can find, by continuity, a small closed ball $\overline{B_\epsilon(x_0, y_0)} \subseteq \mathcal{S}$ such that

$$\mathcal{L}\varphi > 0 \quad \text{on} \quad \overline{B_\epsilon(x_0, y_0)}.$$

Since $v_+ - \varphi$ is u.s.c. and $\overline{B_\epsilon(x_0, y_0)} \setminus B_{\epsilon/2}(x_0, y_0)$ is compact, there exists a $\delta > 0$ such that

$$v_+ - \varphi \leq -\delta \quad \text{on} \quad \overline{B_\epsilon(x_0, y_0)} \setminus B_{\epsilon/2}(x_0, y_0).$$

²In fact, equalities hold for v_+ on the boundary; the reverse inequalities come from the simple fact that (SP1) is preserved under pointwise infimum.

By [9, Proposition 4.1] and Lemma 4.3.5, v_+ can be approximated from above by a non-increasing sequence of stochastic supersolutions v_n . By [11, Lemma 2.4], there exists a large enough N such that $v := v_N$ satisfies

$$v - \varphi \leq -\frac{\delta}{2} \quad \text{on } \overline{B_\epsilon(x_0, y_0)} \setminus B_{\epsilon/2}(x_0, y_0).$$

Choose $\eta \in (0, \delta/2)$ small so that $\varphi^\eta := \varphi - \eta$ satisfies

$$(4.10) \quad \mathcal{L}\varphi^\eta > 0 \quad \text{on } \overline{B_\epsilon(x_0, y_0)}.$$

We also have

$$(4.11) \quad v \leq \varphi - \frac{\delta}{2} < \varphi - \eta = \varphi^\eta \quad \text{on } B_\epsilon(x_0, y_0) \setminus B_{\epsilon/2}(x_0, y_0),$$

and

$$(4.12) \quad \varphi^\eta(x_0, y_0) = \varphi(x_0, y_0) - \eta = v_+(x_0, y_0) - \eta < v_+(x_0, y_0).$$

Define

$$v^\eta := \begin{cases} v \wedge \varphi^\eta & \text{on } \overline{B_\epsilon(x_0, y_0)}, \\ v & \text{on } \overline{B_\epsilon(x_0, y_0)}^c. \end{cases}$$

If we can show $v^\eta \in \mathcal{V}^+$, then (4.12) will lead to a contradiction to the (pointwise) minimality of v_+ . Clearly, v^η is u.s.c. since the minimum of u.s.c. functions is u.s.c. and $v^\eta = v$ outside $B_{\epsilon/2}(x_0, y_0)$. Boundedness is also easy. (SP1) is satisfied because $v^\eta = v$ on $\partial\mathcal{S}$. The remaining proof of case (i) is devoted to the verification of (SP2), i.e. the supermartingale property.

Let (τ, ξ) be any random initial condition and (B^0, S^0) be the (τ, ξ) -admissible control in (SP2) for the stochastic supersolution v . Let

$$A := \{\xi \in B_{\epsilon/2}(x_0, y_0)\} \cap \{\varphi^\eta(\xi) < v(\xi)\} \in \mathcal{G}_\tau.$$

Define a new control

$$(B_t^1, S_t^1) := 1_{A^c \cap \{t \geq \tau\}}(B_t^0 - B_{\tau-}^0, S_t^0 - S_{\tau-}^0).$$

(B^1, S^1) follows (B^0, S^0) starting from time τ when the position ξ satisfies $v^\eta(\xi) = v(\xi)$, i.e. when it is optimal to use the control corresponding to v . By Lemma 4.2.4.i, $(B^1, S^1) \in \mathcal{A}(\tau, \xi)$. Let

$$\tau_1 := \inf\{t \in [\tau, \sigma^{\tau, \xi, B^1, S^1}] : (X_t^{\tau, \xi, B^1, S^1}, Y_t^{\tau, \xi, B^1, S^1}) \notin B_{\epsilon/2}(x_0, y_0)\}$$

be the exit time of the ball $B_{\epsilon/2}(x_0, y_0)$ and

$$\xi_1 := (X_{\tau_1}^{\tau, \xi, B^1, S^1}, Y_{\tau_1}^{\tau, \xi, B^1, S^1}) \in \mathcal{G}_{\tau_1}$$

be the exit position. Since X^{τ, ξ, B^1, S^1} and Y^{τ, ξ, B^1, S^1} are RCLL, we have $\xi_1 \notin B_{\epsilon/2}(x_0, y_0)$.³

By Lemma 4.2.4.ii, (τ_1, ξ_1) is a valid random initial condition. Let (B^2, S^2) be the (τ_1, ξ_1) -admissible control in (SP2) for v . Set

$$(B_t, S_t) := (B_t^1, S_t^1)1_{\{t < \tau_1\}} + (B_t^2 - B_{\tau_1-}^2 + B_{\tau_1}^1, S_t^2 - S_{\tau_1-}^2 + S_{\tau_1}^1)1_{\{t \geq \tau_1\}}.$$

Note that we allow “double transactions” at time τ_1 , first by $(\Delta B_{\tau_1}^1, \Delta S_{\tau_1}^1)$, then by $(\Delta B_{\tau_1}^2, \Delta S_{\tau_1}^2)$. Lemma 4.2.4.ii also implies $(B, S) \in \mathcal{A}(\tau, \xi)$. We now check the supermartingale property (SP2) for v^η with control (B, S) .

Let ρ be any \mathbb{G} -stopping time taking values in $[\tau, \sigma^{\tau, \xi, B, S}]$. In the event A , $(B^1, S^1) = 0$ so that $(X^{\tau, \xi, B^1, S^1}, Y^{\tau, \xi, B^1, S^1})$ exits $B_{\epsilon/2}(x_0, y_0)$ either by diffusion or by death, giving $\xi_1 \in \partial B_{\epsilon/2}(x_0, y_0) \cup \{\Delta\}$. The control (B, S) is inactive before time

³If $\xi \notin B_{\epsilon/2}(x_0, y_0)$, it is possible for the process to immediately jump back to $B_{\epsilon/2}(x_0, y_0)$ at time τ . In this case, although we start outside the ball, $\tau_1 \neq \tau$ because $(X_t^{\tau, \xi, B^1, S^1}, Y_t^{\tau, \xi, B^1, S^1})$ gives the post-jump position at time t which is inside the ball at $t = \tau$, and will stay inside the ball for some positive amount of time by the right continuity of its paths.

τ_1 and equals $(\Delta B_{\tau_1}^2, \Delta S_{\tau_1}^2)$ at τ_1 . By Itô's formula, we have in the event A

$$\begin{aligned}
& \varphi^\eta(X_{\rho \wedge \tau_1}^{\tau, \xi, B, S}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B, S}) - \varphi^\eta(X_\tau^{\tau, \xi, B, S}, Y_\tau^{\tau, \xi, B, S}) \\
&= \int_\tau^{\rho \wedge \tau_1} -\mathcal{L}\varphi^\eta(X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S})dt + \int_\tau^{\rho \wedge \tau_1} (\varphi^\eta)'(X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S})\sigma Y_t^{\tau, \xi, B, S}dW_t \\
&+ \int_\tau^{\rho \wedge \tau_1} \left[\varphi^\eta(\Delta) - \varphi^\eta(X_{t-}^{\tau, \xi, B, S}, Y_{t-}^{\tau, \xi, B, S}) \right] d(N_t - \beta t) \\
&+ \mathbf{1}_{\{\rho \geq \tau_1\}} [\varphi^\eta(\xi_1 + \Delta\xi) - \varphi^\eta(\xi_1)],
\end{aligned}$$

where

$$\Delta\xi := (-1, 1 - \lambda)\Delta B_{\tau_1}^2 + (1 - \mu, -1)\Delta S_{\tau_1}^2.$$

Since $(X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S}) \in B_{\epsilon/2}(x_0, y_0)$ for $\tau \leq t < \tau_1$ on A , and $\mathcal{L}\varphi^\eta > 0$ in $B_{\epsilon/2}(x_0, y_0)$ by (4.10), the dt -integral is non-positive. The integrals with respect to the Brownian motion and the compensated Poisson process vanish by taking \mathcal{G}_τ -conditional expectation. We therefore obtain

$$\begin{aligned}
& \mathbb{E}[1_A \varphi^\eta(X_{\rho \wedge \tau_1}^{\tau, \xi, B, S}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B, S}) - 1_{A \cap \{\rho \geq \tau_1\}} (\varphi^\eta(\xi_1 + \Delta\xi) - \varphi^\eta(\xi_1)) | \mathcal{G}_\tau] \\
&\leq 1_A \varphi^\eta(X_\tau^{\tau, \xi, B, S}, Y_\tau^{\tau, \xi, B, S}) = 1_{A \cap \{\tau < \tau_d\}} \varphi^\eta(\xi) + 1_{A \cap \{\tau = \tau_d\}} \varphi^\eta(\Delta) \\
&= 1_A \varphi^\eta(\xi) \leq 1_A v^\eta(\xi).
\end{aligned}$$

In the last equality, we used $\xi = \Delta$ if $\tau = \tau_d$. Notice that

$$\begin{aligned}
& 1_A \varphi^\eta(X_{\rho \wedge \tau_1}^{\tau, \xi, B, S}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B, S}) - 1_{A \cap \{\rho \geq \tau_1\}} (\varphi^\eta(\xi_1 + \Delta\xi) - \varphi^\eta(\xi_1)) \\
&= 1_{A \cap \{\rho < \tau_1\}} \varphi^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A \cap \{\rho \geq \tau_1\}} \varphi^\eta(\xi_1 + \Delta\xi) \\
&\quad - 1_{A \cap \{\rho \geq \tau_1\}} (\varphi^\eta(\xi_1 + \Delta\xi) - \varphi^\eta(\xi_1)) \\
&= 1_{A \cap \{\rho < \tau_1\}} \varphi^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A \cap \{\rho \geq \tau_1\}} \varphi^\eta(\xi_1).
\end{aligned}$$

So

$$\mathbb{E}[1_{A \cap \{\rho < \tau_1\}} \varphi^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A \cap \{\rho \geq \tau_1\}} \varphi^\eta(\xi_1) | \mathcal{G}_\tau] \leq 1_A v^\eta(\xi).$$

We have argued that $\xi_1 \in \partial B_{\epsilon/2}(x_0, y_0) \cup \{\Delta\}$ on A . By (4.11) and the definition of v^η , we know $v^\eta \leq \varphi^\eta$ in $B_\epsilon(x_0, y_0)$. This allows us to replace φ^η by v^η in the above inequality and get

$$(4.13) \quad \mathbb{E}[1_{A \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A \cap \{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \leq 1_A v^\eta(\xi).$$

By “optimality” of (B^0, S^0) (and thus (B^1, S^1) on A^c) for v with random initial condition (τ, ξ) , we have

$$\mathbb{E}[1_{A^c} v(X_{\rho \wedge \tau_1}^{\tau, \xi, B^1, S^1}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B^1, S^1}) | \mathcal{G}_\tau] \leq 1_{A^c} v(\xi).$$

Since $v^\eta \leq v$ everywhere, we can replace v by v^η on the left hand side in the above inequality. On A^c , either $\xi \notin B_{\epsilon/2}(x_0, y_0)$, or $\xi \in B_{\epsilon/2}(x_0, y_0)$ and $v(\xi) \leq \varphi^\eta(\xi)$. In both cases, $v(\xi) = v^\eta(\xi)$ since $v^\eta = v$ outside the ball $B_{\epsilon/2}(x_0, y_0)$. So we can also replace v by v^η on the right hand side. Splitting the set A^c on the left hand side according to the relation between ρ and τ_1 , and using the definition of (B, S) , we have

$$(4.14) \quad \mathbb{E}[1_{A^c \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A^c \cap \{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \leq 1_{A^c} v^\eta(\xi).$$

Combining (4.13) and (4.14) gives us

$$(4.15) \quad \mathbb{E}[1_{\{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \leq v^\eta(\xi).$$

By “optimality” of (B^2, S^2) for v with random initial condition (τ_1, ξ_1) , we have (by applying the supermartingale property to the stopping time $\rho \vee \tau_1$)

$$\mathbb{E}[1_{\{\rho \geq \tau_1\}} v(X_\rho^{\tau_1, \xi_1, B^2, S^2}, Y_\rho^{\tau_1, \xi_1, B^2, S^2}) | \mathcal{G}_{\tau_1}] \leq 1_{\{\rho \geq \tau_1\}} v(\xi_1).$$

Same as before, we can replace all v 's by v^η in the above inequality because $v^\eta \leq v$ everywhere, $v = v^\eta$ outside $B_{\epsilon/2}(x_0, y_0)$ and ξ_1 , being the exit position, is outside

$B_{\epsilon/2}(x_0, y_0)$. So

$$\begin{aligned}\mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_{\tau_1}] &= \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau_1, \xi_1, B^2, S^2}, Y_\rho^{\tau_1, \xi_1, B^2, S^2}) | \mathcal{G}_{\tau_1}] \\ &\leq 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1).\end{aligned}$$

Taking \mathcal{G}_τ -condition expectation and using tower property yields

$$(4.16) \quad \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) - 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \leq 0.$$

Finally, we add (4.15) and (4.16) to get

$$\mathbb{E}[v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \leq v^\eta(\xi).$$

This completes the proof of (SP2) and hence of case (i).

Case (ii). $-(1 - \mu)\varphi_x(x_0, y_0) + \varphi_y(x_0, y_0) > 0$. The proof is in most part similar to that of case (i). So we shall be brief on the similar parts. Same as in case (i), we can find $\epsilon, \eta > 0$ and $v \in \mathcal{V}^+$ such that $\varphi^\eta := \varphi - \eta$ satisfies

$$(4.17) \quad \begin{aligned}-(1 - \mu)\varphi_x^\eta + \varphi_y^\eta &> 0 \quad \text{on} \quad \overline{B_\epsilon(x_0, y_0)}, \\ v &\leq \varphi^\eta \quad \text{on} \quad \overline{B_\epsilon(x_0, y_0)} \setminus B_{\epsilon/2}(x_0, y_0), \\ \varphi^\eta(x_0, y_0) &< v_+(x_0, y_0).\end{aligned}$$

Define

$$v^\eta := \begin{cases} v \wedge \varphi^\eta & \text{on } \overline{B_\epsilon(x_0, y_0)}, \\ v & \text{on } \overline{B_\epsilon(x_0, y_0)}^c. \end{cases}$$

It suffices to show $v^\eta \in \mathcal{V}^+$. And the only nontrivial part is to check v^η satisfies (SP2).

Let (τ, ξ) be any random initial condition and (B^0, S^0) be a (τ, ξ) -admissible control in (SP2) for the stochastic supersolution v . Let

$$A := \{\xi \in B_{\epsilon/2}(x_0, y_0)\} \cap \{\varphi^\eta(\xi) < v(\xi)\} \in \mathcal{G}_\tau.$$

Observe that (4.17) implies for any $(x, y) \in B_\epsilon(x_0, y_0)$ and $h > 0$ small such that $(x + (1 - \mu)h, y - h) \in B_\epsilon(x_0, y_0)$, we have

$$(4.18) \quad \varphi^\eta(x + (1 - \mu)h, y - h) - \varphi^\eta(x, y) = h[(1 - \mu)\varphi_x^\eta - \varphi_y^\eta](x + (1 - \mu)h, y - h) < 0$$

for some $h' \in (0, h)$ by Mean Value Theorem. This suggests selling stocks is optimal on the set A . Given a point $(x, y) \in B_{\epsilon/2}(x_0, y_0)$, denote by $\mathfrak{s}(x, y) = (\mathfrak{s}^0(x, y), \mathfrak{s}^1(x, y))$ the intersection of the ray $\{(x + (1 - \mu)h, y - h) : h \geq 0\}$ and $\partial B_{\epsilon/2}(x_0, y_0)$, i.e. the unique point on $\partial B_{\epsilon/2}(x_0, y_0)$ that can be reached by a sell. Define a new control

$$(B_t^1, S_t^1) := 1_{A \cap \{t \geq \tau\}}(0, \xi^1 - \mathfrak{s}^1(\xi)) + 1_{A^c \cap \{t \geq \tau\}}(B_t^0 - B_{\tau-}^0, S_t^0 - S_{\tau-}^0).$$

(B^1, S^1) says starting at time τ , if we are in A , we immediately jump to $\partial B_{\epsilon/2}(x_0, y_0)$ by a sell and do nothing afterwards; if we are in A^c , we follow (B^0, S^0) . A slight variation of Lemma 4.2.4.i shows (B^1, S^1) is (τ, ξ) -admissible. Let

$$\tau_1 := \inf\{t \in [\tau, \sigma^{\tau, \xi, B^1, S^1}] : (X_t^{\tau, \xi, B^1, S^1}, Y_t^{\tau, \xi, B^1, S^1}) \notin B_{\epsilon/2}(x_0, y_0)\}$$

be the exit time of the ball $B_{\epsilon/2}(x_0, y_0)$ and

$$\xi_1 := (X_{\tau_1}^{\tau, \xi, B^1, S^1}, Y_{\tau_1}^{\tau, \xi, B^1, S^1}) \in \mathcal{G}_{\tau_1}$$

be the exit position. As in case (i), $\xi_1 \notin B_{\epsilon/2}(x_0, y_0)$ and (τ_1, ξ_1) is a valid random initial condition. Also notice that on A , $\tau_1 = \tau$ and $\xi_1 = \mathfrak{s}(\xi)$ if $\tau < \tau_d$. Let (B^2, S^2) be a (τ_1, ξ_1) -admissible control in (SP2) for v . Set

$$(B_t, S_t) := (B_t^1, S_t^1)1_{\{t < \tau_1\}} + (B_t^2 - B_{\tau_1-}^2 + B_{\tau_1}^1, S_t^2 - S_{\tau_1-}^2 + S_{\tau_1}^1)1_{\{t \geq \tau_1\}}.$$

$(B, S) \in \mathcal{A}(\tau, \xi)$ by Lemma 4.2.4.ii. It remains to check (SP2) for v^η with control (B, S) .

Let ρ be any \mathbb{G} -stopping time taking values in $[\tau, \sigma^{\tau, \xi, B, S}]$. In the event A (recall that $\tau_1 = \tau$), when $\tau < \tau_d$, (4.18) implies $\varphi^\eta(\xi_1) = \varphi^\eta(\mathfrak{s}(\xi)) < \varphi^\eta(\xi)$; when $\tau = \tau_d$, $\varphi^\eta(\xi_1) = \varphi^\eta(\xi) = \varphi^\eta(\Delta) = 0$. So

$$(4.19) \quad 1_A v^\eta(\xi_1) \leq 1_A \varphi^\eta(\xi_1) < 1_A \varphi^\eta(\xi) = 1_A v^\eta(\xi).$$

In the event A^c , we use that (B^0, S^0) is “optimal” for v to obtain

$$(4.20) \quad \begin{aligned} & \mathbb{E}[1_{A^c \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B^1, S^1}, Y_\rho^{\tau, \xi, B^1, S^1}) + 1_{A^c \cap \{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \\ &= \mathbb{E}[1_{A^c} v^\eta(X_{\rho \wedge \tau_1}^{\tau, \xi, B^1, S^1}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B^1, S^1}) | \mathcal{G}_\tau] \leq \mathbb{E}[1_{A^c} v(X_{\rho \wedge \tau_1}^{\tau, \xi, B^1, S^1}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B^1, S^1}) | \mathcal{G}_\tau] \\ &= \mathbb{E}[1_{A^c} v(X_{\rho \wedge \tau_1}^{\tau, \xi, B^0, S^0}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B^0, S^0}) | \mathcal{G}_\tau] \leq 1_{A^c} v(\xi) = 1_{A^c} v^\eta(\xi). \end{aligned}$$

Combining (4.19) and (4.20), and using that (B, S) equals (B^1, S^1) on $[\tau, \tau_1]$, we get

$$(4.21) \quad \mathbb{E}[1_{\{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \leq v^\eta(\xi).$$

By “optimality” of (B^2, S^2) for v with random initial condition (τ_1, ξ_1) , we have

$$\begin{aligned} \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_{\tau_1}] &= \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau_1, \xi_1, B^2, S^2}, Y_\rho^{\tau_1, \xi_1, B^2, S^2}) | \mathcal{G}_{\tau_1}] \\ &\leq \mathbb{E}[1_{\{\rho \geq \tau_1\}} v(X_\rho^{\tau_1, \xi_1, B^2, S^2}, Y_\rho^{\tau_1, \xi_1, B^2, S^2}) | \mathcal{G}_{\tau_1}] \\ &\leq 1_{\{\rho \geq \tau_1\}} v(\xi_1) = 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1). \end{aligned}$$

Taking \mathcal{G}_τ -condition expectation yields

$$(4.22) \quad \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) - 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \leq 0.$$

Finally, we add (4.21) and (4.22) to get

$$\mathbb{E}[v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \leq v^\eta(\xi).$$

This completes the proof of case (ii).

Case (iii). $\varphi_x(x_0, y_0) - (1 - \lambda)\varphi_y(x_0, y_0) > 0$. This case is symmetric to case (ii). Buying stock is optimal in a neighborhood of (x_0, y_0) . We define the set A

and the “optimal” (τ, ξ) -admissible control in the same way as in case (ii) except one modification: in the definition of (B^1, S^1) , $(0, \xi^1 - \mathfrak{s}^1(\xi))$ is replaced by $(\xi^0 - \mathfrak{b}^0(\xi), 0)$, where for $(x, y) \in B_{\epsilon/2}(x, y)$, $\mathfrak{b}(x, y)$ is defined to be the intersection of the ray $\{x - h, y + (1 - \lambda)h : h \geq 0\}$ and $\partial B_{\epsilon/2}(x_0, y_0)$, i.e. the unique point on $\partial B_{\epsilon/2}(x_0, y_0)$ that can be reached by a buy. The rest of the argument is almost the same. \square

4.4 Stochastic subsolution

Definition 4.4.1. A bounded l.s.c. function v on $\bar{\mathcal{S}}$ is called a *stochastic subsolution* of (4.4), (4.5) if

(SB1) $v \leq 1$ on $\partial \mathcal{S}_b$, $v \leq 0$ on $\partial \mathcal{S}_{c/\tau}$;

(SB2) for any random initial condition (τ, ξ) , control pair $(B, S) \in \mathcal{A}(\tau, \xi)$ and \mathbb{G} -stopping time $\rho \in [\tau, \sigma^{\tau, \xi, B, S}]$,

$$\mathbb{E}[v(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \geq v(\xi),$$

where v is understood to be its extension to $\bar{\mathcal{S}} \cup \{\Delta\}$.

Denote the set of stochastic subsolutions by \mathcal{V}^- .

Remark 4.4.2. $\mathcal{V}^- \neq \emptyset$ since the constant $0 \in \mathcal{V}^-$. Similar to the stochastic supersolution case, there is also a member of \mathcal{V}^- which satisfies (SB1) with equalities, namely, the lower bound function $\underline{\psi}$ defined in (4.9). (See Lemma 4.4.6.)

Remark 4.4.3. Any stochastic subsolution v is dominated by the value function ψ on $\bar{\mathcal{S}}$. Indeed, on $\partial \mathcal{S}$, we clearly have $v \leq \psi$ by (SP1). For $(x, y) \in \mathcal{S}$, take $\tau = 0$, $\xi = (x, y)$, (B, S) be any (x, y) -admissible control, and $\rho = \sigma^{x, y, B, S}$. We have by

(SB2) and (SB1) that

$$\begin{aligned} v(x, y) &\leq \mathbb{E}[v(X_\rho^{x,y,B,S}, Y_\rho^{x,y,B,S})] \leq \mathbb{E}\left[1_{\{\rho=\tau_b^{x,y,B,S}\}}\right] \\ &= \mathbb{P}(\tau_b^{x,y,B,S} < \tau_d \wedge \tau_s^{x,y,B,S}) \leq \mathbb{P}(\tau_b^{x,y,B,S} < \tau_d). \end{aligned}$$

Since this holds for any $(B, S) \in \mathcal{A}(x, y)$, taking infimum yields

$$v(x, y) \leq \inf_{(B,S) \in \mathcal{A}(x,y)} \mathbb{P}(\tau_b^{x,y,B,S} < \tau_d) = \psi(x, y).$$

Lemma 4.4.4. *Let $v_1, v_2 \in \mathcal{V}^-$. Then $v_1 \vee v_2 \in \mathcal{V}^-$.*

Proof. The maximum of bounded l.s.c. functions is still bounded and l.s.c.. (SB1) is clearly stable under maximum. For (SB2), simply notice that

$$\mathbb{E}[(v_1 \vee v_2)(X_\rho^{\tau,\xi,B,S}, Y_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] \geq \mathbb{E}[v_i(X_\rho^{\tau,\xi,B,S}, Y_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] \geq v_i(\xi), \quad i = 1, 2.$$

So

$$\mathbb{E}[(v_1 \vee v_2)(X_\rho^{\tau,\xi,B,S}, Y_\rho^{\tau,\xi,B,S})|\mathcal{G}_\tau] \geq (v_1 \vee v_2)(\xi).$$

□

Remark 4.4.5. The above proof can be easily generalized to the countable case. In particular, the supremum of a countable family of stochastic subsolutions is bounded from above because every stochastic subsolution is dominated by the value function. In fact, it also generalizes to the uncountable case by [9, Proposition 4.1] which says the supremum of an uncountable family of l.s.c. functions equals the supremum over some countable subfamily.

Lemma 4.4.6. $\underline{\psi} \in \mathcal{V}^-$.

Proof. Recall that $\underline{\psi}$ can be written as the supremum of all ψ_k 's with $k \in (1 - \mu, \frac{1}{1-\lambda}) \cap \mathbb{Q}$ where ψ_k is defined in (4.7). To show $\underline{\psi} \in \mathcal{V}^-$, it suffices to show $\psi_k \in \mathcal{V}^-$

for $k \in (1 - \mu, \frac{1}{1-\lambda})$ by Lemma 4.4.4 and the remark after it. To see (SB2) holds for ψ_k , let (τ, ξ) be any random initial condition, (B, S) be any (τ, ξ) -admissible control and $\rho \in [\tau, \sigma^{\tau, \xi, B, S}]$ be any \mathbb{G} -stopping time. For brevity, we shall omit the superscripts (τ, ξ, B, S) in all controlled processes and relevant stopping times in the rest of this proof. For functions defined on $[b, \infty)$, we extend them to $[b, \infty) \cup \{\Delta\}$ by assigning zero to the function value at Δ . Define a new process

$$Z_t := \begin{cases} X_t + kY_t, & t < \tau_d, \\ \Delta, & t \geq \tau_d. \end{cases}$$

Observe that $Z_t \in [b, \infty) \cup \{\Delta\}$ for all $t \in [\tau, \rho]$. We also have

$$dZ_t = (rZ_t + (\alpha - r)kY_t - c)dt + \sigma kY_t dW_t + [k(1 - \lambda) - 1]dB_t + (1 - \mu - k)dS_t, \quad t < \tau_d.$$

Since $1 - \mu < k < \frac{1}{1-\lambda}$, the dB and dS terms are non-positive. So for $t \geq \tau$, Z_t is bounded above by the process \tilde{Z}_t defined by

$$d\tilde{Z}_t = (r\tilde{Z}_t + (\alpha - r)kY_t - c)dt + \sigma kY_t dW_t, \quad \tilde{Z}_\tau = \xi^0 + k\xi^1 \quad \text{for } t < \tau_d,$$

and $\tilde{Z}_t = \Delta$ for $t \geq \tau_d$. \tilde{Z}_t is the wealth process if the amount invested in the (frictionless) stock market is kY_t . Let $f(x) := \psi_k(x, 0) \in C^1[b, c/r] \cap C^2[b, c/r]$. We have $\psi_k(x, y) = f(x + ky)$. Since f is decreasing in $[b, \infty)$, we deduce

$$(4.23) \quad \mathbb{E}[\psi_k(X_\rho, Y_\rho) | \mathcal{G}_\tau] = \mathbb{E}[1_{\{\rho < \tau_d\}} f(Z_\rho) | \mathcal{G}_\tau] \geq \mathbb{E}[1_{\{\rho < \tau_d\}} f(\tilde{Z}_\rho) | \mathcal{G}_\tau] = \mathbb{E}[f(\tilde{Z}_\rho) | \mathcal{G}_\tau].$$

In the event $A := \{\tilde{Z}_\tau \in [c/r, \infty) \cup \{\Delta\}\} \in \mathcal{G}_\tau$, we have $f(\tilde{Z}_\rho) \geq 0 = f(\tilde{Z}_\tau)$. In the event $A^c := \{\tilde{Z}_\tau \in [b, c/r)\}$, we let $\nu := \inf\{t \geq 0 : \tilde{Z}_\rho \in [c/r, \infty)\}$, and use f is non-negative in $[b, \infty)$ and zero in $[c/r, \infty)$ to get $f(\tilde{Z}_\rho) \geq f(\tilde{Z}_{\rho \wedge \nu})$. We therefore have

$$(4.24) \quad \mathbb{E}[f(\tilde{Z}_\rho) | \mathcal{G}_\tau] \geq \mathbb{E}[1_A f(\tilde{Z}_\tau) + 1_{A^c} f(\tilde{Z}_{\rho \wedge \nu}) | \mathcal{G}_\tau].$$

In the event A^c , we use Itô's formula to obtain

$$\begin{aligned} & f(\tilde{Z}_{\rho \wedge \nu}) - f(\tilde{Z}_\tau) \\ &= \int_\tau^{\rho \wedge \nu} \left\{ f'(\tilde{Z}_t)[r\tilde{Z}_t + (\alpha - r)kY_t - c] + \frac{1}{2}f''(\tilde{Z}_t)\sigma^2(kY_t)^2 - \beta f(\tilde{Z}_t) \right\} dt \\ & \quad + \int_\tau^{\rho \wedge \nu} f'(\tilde{Z}_t)\sigma kY_t dW_t + \int_\tau^{\rho \wedge \nu} [f(\Delta) - f(\tilde{Z}_{t-})]d(N_t - \beta t). \end{aligned}$$

Notice that f is the frictionless value function which satisfies the HJB equation

$$\beta f(x) = \inf_\pi \left\{ \frac{1}{2}f''(x)\pi^2 + (\alpha - r)f'(x)\pi + (rx - c)f'(x) \right\}$$

in $[b, c/r]$. It follows that the drift term is non-negative. For $t \in [\tau, \rho \wedge \nu]$, the process $\tilde{Z}_t \in [b, c/r]$. So $Z_t \in [b, c/r]$ and the process (X_t, Y_t) stays inside the bounded set $\{(x, y) \in \bar{\mathcal{S}} : x + ky \leq c/r\}$. Here it is crucial that $k \in (1 - \mu, \frac{1}{1-\lambda})$ for Y_t to be bounded. The integrals with respect to the martingales W_t and $N_t - \beta t$ then vanish upon taking \mathcal{G}_τ -conditional expectation. This leads to

$$(4.25) \quad \mathbb{E}[1_{A^c} f(\tilde{Z}_{\rho \wedge \nu}) | \mathcal{G}_\tau] \geq \mathbb{E}[1_{A^c} f(\tilde{Z}_\tau) | \mathcal{G}_\tau].$$

Putting (4.23), (4.24) and (4.25) together, we get

$$\mathbb{E}[\psi_k(X_\rho, Y_\rho) | \mathcal{G}_\tau] \geq f(\tilde{Z}_\tau) = 1_{\{\tau < \tau_d\}} f(\xi^0 + k\xi^1) = 1_{\{\tau < \tau_d\}} \psi_k(\xi) = \psi_k(\xi)$$

which is the desired submartingale property. \square

Proposition 4.4.7. *The lower stochastic envelope*

$$v_-(x, y) := \sup_{v \in \mathcal{V}^-} v(x, y)$$

is a viscosity supersolution of (4.4) satisfying $v_- \geq 1$ on $\partial \mathcal{S}_b$ and $v_- \geq 0$ on $\partial \mathcal{S}_{c/r}$.

Proof. The boundary inequalities are satisfied because $v_- \geq \underline{\psi}$ by Lemma 4.4.6.⁴ To show interior viscosity supersolution property, let $(x_0, y_0) \in \mathcal{S}$ and $\varphi \in C^2(\mathcal{S})$ be a

⁴In fact, equalities hold for v_- on the boundary; the reverse inequalities holds because (SB1) is preserved under pointwise maximum.

test function such that $v_- - \varphi$ attains a strict minimum of zero at (x_0, y_0) . We need to show

$$\max \{ \mathcal{L}\varphi, -(1-\mu)\varphi_x + \varphi_y, \varphi_x - (1-\lambda)\varphi_y \} (x_0, y_0) \geq 0.$$

Assume on the contrary that

$$\max \{ \mathcal{L}\varphi, -(1-\mu)\varphi_x + \varphi_y, \varphi_x - (1-\lambda)\varphi_y \} (x_0, y_0) < 0.$$

Similar to the proof of Proposition 4.3.6, we can find $0 < \epsilon < 1$, $\eta > 0$ and $v \in \mathcal{V}^-$ such that $\varphi^\eta := \varphi + \eta$ satisfies

$$(4.26) \quad \max \{ \mathcal{L}\varphi^\eta, -(1-\mu)\varphi_x^\eta + \varphi_y^\eta, \varphi_x^\eta - (1-\lambda)\varphi_y^\eta \} < 0 \quad \text{on} \quad \overline{B_\epsilon(x_0, y_0)},$$

$$(4.27) \quad \varphi^\eta \leq v \quad \text{on} \quad \overline{B_\epsilon(x_0, y_0)} \setminus B_{\epsilon/2}(x_0, y_0),$$

and

$$\varphi^\eta(x_0, y_0) > v_-(x_0, y_0).$$

The technique for constructing the lifting function φ^η is classical and similar to the stochastic supersolution case. So we skip the details. Define

$$v^\eta := \begin{cases} v \vee \varphi^\eta & \text{on } \overline{B_\epsilon(x_0, y_0)}, \\ v & \text{on } \overline{B_\epsilon(x_0, y_0)}^c. \end{cases}$$

It suffices to show $v^\eta \in \mathcal{V}^-$. And the only nontrivial part is to check v^η satisfies (SB2).

Let (τ, ξ) be any random initial condition, (B, S) be any (τ, ξ) -admissible control and $\rho \in [\tau, \sigma^{\tau, \xi, B, S}]$ be any \mathbb{G} -stopping time. Let

$$A := \{ \xi \in B_{\epsilon/2}(x_0, y_0) \} \cap \{ \varphi^\eta(\xi) > v(\xi) \} \in \mathcal{G}_\tau.$$

Let

$$\tau_1 := \inf\{t \in [\tau, \sigma^{\tau, \xi, B, S}] : (X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S}) \notin B_{\epsilon/2}(x_0, y_0)\}$$

and

$$\xi_1 := (X_{\tau_1}^{\tau, \xi, B, S}, Y_{\tau_1}^{\tau, \xi, B, S}) \in \mathcal{G}_{\tau_1}.$$

In the event A , because of a possible jump transaction at time τ_1 , ξ_1 may not be on $\partial B_{\epsilon/2}(x_0, y_0) \cup \{\Delta\}$. This will bring some problem since (4.26) is only valid locally. To overcome this issue, we define an intermediate position ξ'_1 as follows: let $\xi_{1-} := (X_{\tau_1-}^{\tau, \xi, B, S}, Y_{\tau_1-}^{\tau, \xi, B, S})$. We have $\xi_{1-} \in \overline{B_{\epsilon/2}(x_0, y_0)}$ on A . Define

$$\xi'_1 := 1_{A \cap \{\tau_1 < \tau_d\}} (1_{\{\Delta B_{\tau_1} > 0\}} \mathbf{b}(\xi_{1-}) + 1_{\{\Delta S_{\tau_1} > 0\}} \mathbf{s}(\xi_{1-})) + 1_{A^c \cup \{\tau_1 = \tau_d\}} \xi_1 \in \mathcal{G}_{\tau_1},$$

where \mathbf{b}, \mathbf{s} are the functions introduced in cases (ii) and (iii) of the proof of Proposition 4.3.6. On $A \cap \{\tau_1 < \tau_d\}$, ξ'_1 is the intersection of $\partial B_{\epsilon/2}(x_0, y_0)$ and the line segment connecting ξ_{1-} and ξ_1 . Also define (B^1, S^1) by

$$\begin{aligned} (\Delta B_{\tau_1}^1, \Delta S_{\tau_1}^1) &:= 1_{A \cap \{\tau_1 < \tau_d\}} (1_{\{\Delta B_{\tau_1} > 0\}} (\xi_{1-}^0 - \mathbf{b}^0(\xi_{1-}), 0) + 1_{\{\Delta S_{\tau_1} > 0\}} (0, \xi_{1-}^1 - \mathbf{s}^1(\xi_{1-}))) \\ &\quad + 1_{A^c \cup \{\tau_1 = \tau_d\}} (\Delta B_{\tau_1}, \Delta S_{\tau_1}) \end{aligned}$$

and

$$(B_t^1, S_t^1) := 1_{\{t < \tau_1\}} (B_t, S_t) + 1_{\{t \geq \tau_1\}} [(B_{\tau_1-}, S_{\tau_1-}) + (\Delta B_{\tau_1}^1, \Delta S_{\tau_1}^1)].$$

That is, (B^1, S^1) agrees with (B, S) before time τ_1 , but at time τ_1 , the corresponding controlled process only jumps to ξ'_1 instead of ξ_1 . We have $(B^1, S^1) \in \mathcal{A}_0$ and $(X_t^{\tau, \xi, B^1, S^1}, Y_t^{\tau, \xi, B^1, S^1}) \in \overline{B_{\epsilon/2}(x_0, y_0)} \cup \{\Delta\}$ for all $t \in [\tau, \tau_1]$ on A . Apply generalized

Itô's formula to the RCLL semimartingale $\varphi^\eta(X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1})$ on A , we get

$$\begin{aligned}
& \varphi^\eta(X_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}, Y_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}) - \varphi^\eta(X_\tau^{\tau,\xi,B^1,S^1}, Y_\tau^{\tau,\xi,B^1,S^1}) \\
&= \int_\tau^{\rho \wedge \tau_1} -\mathcal{L}\varphi^\eta(X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) dt \\
&+ \int_\tau^{\rho \wedge \tau_1} (\varphi^\eta)'(X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) \sigma Y_t^{\tau,\xi,B^1,S^1} dW_t \\
&+ \int_\tau^{\rho \wedge \tau_1} [-\varphi_x^\eta + (1-\lambda)\varphi_y^\eta](X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) dB_t^c \\
&+ \int_\tau^{\rho \wedge \tau_1} [(1-\mu)\varphi_x^\eta - \varphi_y^\eta](X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) dS_t^c \\
&+ \int_\tau^{\rho \wedge \tau_1} [\varphi^\eta(\Delta) - \varphi^\eta(X_{t-}^{\tau,\xi,B^1,S^1}, Y_{t-}^{\tau,\xi,B^1,S^1})] d(N_t - \beta t) \\
&+ \sum_{\substack{\tau \leq t \leq \rho \wedge \tau_1 \\ t < \tau_d}} \varphi^\eta(X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) - \varphi^\eta(X_{t-}^{\tau,\xi,B^1,S^1}, Y_{t-}^{\tau,\xi,B^1,S^1})
\end{aligned}$$

where B^c, S^c denote the continuous part of B, S . By (4.26), the dt , dB^c and dS^c integrals are non-negative. The dW integral and the integral with respect to the compensated Poisson process vanish if we take \mathcal{G}_τ -conditional expectation. We now analyze the last term which represents contribution from jump transactions. Similar to case (ii) of the proof of Proposition 4.3.6 (see (4.18)), we can use (4.26) and Mean Value Theorem to deduce

$$\varphi^\eta(x-h, y+(1-\lambda)h) \geq \varphi^\eta(x, y),$$

and

$$\varphi^\eta(x+(1-\mu)h', y-h') \geq \varphi^\eta(x, y).$$

for all $(x, y) \in B_\epsilon(x_0, y_0)$ and $h, h' > 0$ such that $(x-h, y+(1-\lambda)h), (x+(1-\mu)h', y-h') \in B_\epsilon(x_0, y_0)$. It follows that on the set A and for $t \in [\tau, \tau_1] \setminus \{\tau_d\}$, if $\Delta B_t^1 > 0$, then

$$\begin{aligned}
\varphi^\eta(X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) &= \varphi^\eta(X_{t-}^{\tau,\xi,B^1,S^1} - \Delta B_t^1, Y_{t-}^{\tau,\xi,B^1,S^1} + (1-\lambda)\Delta B_t^1) \\
&\geq \varphi^\eta(X_{t-}^{\tau,\xi,B^1,S^1}, Y_{t-}^{\tau,\xi,B^1,S^1}).
\end{aligned}$$

If $\Delta S_t^1 > 0$, then

$$\begin{aligned}\varphi^\eta(X_t^{\tau,\xi,B^1,S^1}, Y_t^{\tau,\xi,B^1,S^1}) &= \varphi^\eta(X_{t-}^{\tau,\xi,B^1,S^1} + (1-\mu)\Delta S_t^1, Y_{t-}^{\tau,\xi,B^1,S^1} - \Delta S_t^1) \\ &\geq \varphi^\eta(X_{t-}^{\tau,\xi,B^1,S^1}, Y_{t-}^{\tau,\xi,B^1,S^1}).\end{aligned}$$

Since ΔB_t^1 and ΔS_t^1 are not positive at the same time (see the definition of \mathcal{A}_0), each summand in the last term is non-negative. Putting everything together, we obtain by taking \mathcal{G}_τ -conditional expectation of the expression given by Itô's formula that

$$\mathbb{E}[1_A \varphi^\eta(X_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}, Y_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}) | \mathcal{G}_\tau] \geq 1_A \varphi^\eta(X_\tau^{\tau,\xi,B^1,S^1}, Y_\tau^{\tau,\xi,B^1,S^1}).$$

Again, we use that φ^η is non-decreasing if we move northwest along the vector $(-1, 1 - \lambda)$ and southeast along the vector $(1 - \mu, -1)$ inside the ball $B_\epsilon(x_0, y_0)$ to bound the right hand side from below by

$$1_{A \cap \{\tau < \tau_d\}} \varphi^\eta(\xi) + 1_{A \cap \{\tau = \tau_d\}} \varphi^\eta(\Delta) = 1_A \varphi^\eta(\xi) = 1_A v^\eta(\xi).$$

For the left hand side, we use $v^\eta \geq \varphi^\eta$ in $B_\epsilon(x_0, y_0)$ and that $(B^1, S^1) = (B, S)$ before τ_1 to obtain

$$\begin{aligned}1_A \varphi^\eta(X_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}, Y_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}) &\leq 1_A v^\eta(X_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}, Y_{\rho \wedge \tau_1}^{\tau,\xi,B^1,S^1}) \\ &= 1_{A \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau,\xi,B,S}, Y_\rho^{\tau,\xi,B,S}) + 1_{A \cap \{\rho \geq \tau_1\}} v^\eta(\xi'_1).\end{aligned}$$

Hence

$$(4.28) \quad \mathbb{E}[1_{A \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau,\xi,B,S}, Y_\rho^{\tau,\xi,B,S}) + 1_{A \cap \{\rho \geq \tau_1\}} v^\eta(\xi'_1) | \mathcal{G}_\tau] \geq 1_A v^\eta(\xi).$$

Define

$$(B_t^2, S_t^2) := 1_{\{t \geq \tau_1\}} [(B_{\tau_1}, S_{\tau_1}) - (B_{\tau_1}^1, S_{\tau_1}^1)].$$

Starting with the random initial condition (τ_1, ξ'_1) , (B^2, S^2) immediately brings the state process from ξ'_1 back to ξ_1 and stays inactive afterwards. It is easy to see that

$(X^{\tau_1, \xi'_1, B^2, S^2}, Y^{\tau_1, \xi'_1, B^2, S^2})$ either exit \mathcal{S} at time τ_1 with exit position ξ_1 , or at a later time when the control is inactive so that the exit is caused by diffusion or death. In both cases, the exit position belongs to $\partial\mathcal{S} \cup \{\Delta\}$. So $(B^2, S^2) \in \mathcal{A}(\tau_1, \xi'_1)$. Using the submartingale property of $v(X^{\tau_1, \xi'_1, B^2, S^2}, Y^{\tau_1, \xi'_1, B^2, S^2})$, we have

$$v^\eta(\xi_1) = v(\xi_1) = \mathbb{E}[v(X^{\tau_1, \xi'_1, B^2, S^2}, Y^{\tau_1, \xi'_1, B^2, S^2}) | \mathcal{G}_{\tau_1}] \geq v(\xi'_1) = v^\eta(\xi'_1),$$

where the first and the last equalities hold because $\xi_1, \xi'_1 \notin B_{\epsilon/2}(x_0, y_0)$. (4.28) then implies

$$(4.29) \quad \mathbb{E}[1_{A \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A \cap \{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \geq 1_A v^\eta(\xi).$$

On the set A^c , we use the submartingale property (SB2) of $v(X^{\tau, \xi, B, S}, Y^{\tau, \xi, B, S})$ to get

$$\mathbb{E}[1_{A^c} v^\eta(X_{\rho \wedge \tau_1}^{\tau, \xi, B, S}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \geq \mathbb{E}[1_{A^c} v(X_{\rho \wedge \tau_1}^{\tau, \xi, B, S}, Y_{\rho \wedge \tau_1}^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \geq 1_{A^c} v(\xi) = 1_{A^c} v^\eta(\xi),$$

or

$$(4.30) \quad \mathbb{E}[1_{A^c \cap \{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{A^c \cap \{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \geq 1_{A^c} v^\eta(\xi).$$

Adding (4.29) and (4.30) yields

$$(4.31) \quad \mathbb{E}[1_{\{\rho < \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) + 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau] \geq v^\eta(\xi).$$

Let

$$(B_t^3, S_t^3) := (B_t, S_t) - 1_{\{t \geq \tau_1\}} (\Delta B_{\tau_1}, \Delta S_{\tau_1})$$

be the same control as (B, S) , but with any jump transaction at time τ_1 removed.

We have

$$(4.32) \quad (X_t^{\tau, \xi, B, S}, Y_t^{\tau, \xi, B, S}) = (X_t^{\tau_1, \xi_1, B^3, S^3}, Y_t^{\tau_1, \xi_1, B^3, S^3}) \quad \forall t \geq \tau_1.$$

The reason for introducing another control is because our random initial condition allows a jump at initial time. Since ξ_1 already includes the possible jump transactions specified by (B, S) at time τ_1 , we want to avoid doing the same transaction again when using (τ_1, ξ_1) as the new random initial condition. That is, (B^3, S^3) is defined to make (4.32) hold. To see $(B^3, S^3) \in \mathcal{A}(\tau_1, \xi_1)$, first notice that $\sigma^{\tau, \xi, B, S} \geq \tau_1$ by the definition of τ_1 . (4.32) then implies $\sigma^{\tau_1, \xi_1, B^3, S^3} = \sigma^{\tau, \xi, B, S}$. Thus,

$$(X_{\sigma^{\tau_1, \xi_1, B^3, S^3}}^{\tau_1, \xi_1, B^3, S^3}, Y_{\sigma^{\tau_1, \xi_1, B^3, S^3}}^{\tau_1, \xi_1, B^3, S^3}) = (X_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}, Y_{\sigma^{\tau, \xi, B, S}}^{\tau, \xi, B, S}) \in \partial\mathcal{S} \cup \{\Delta\}$$

by the (τ, ξ) -admissibility of (B, S) . The submartingale property (SB2) of $v(X_{\sigma^{\tau_1, \xi_1, B^3, S^3}}^{\tau_1, \xi_1, B^3, S^3}, Y_{\sigma^{\tau_1, \xi_1, B^3, S^3}}^{\tau_1, \xi_1, B^3, S^3})$ (applied to the stopping time $\rho \vee \tau_1$) implies

$$\begin{aligned} \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_{\tau_1}] &= \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau_1, \xi_1, B^3, S^3}, Y_\rho^{\tau_1, \xi_1, B^3, S^3}) | \mathcal{G}_{\tau_1}] \\ &\geq \mathbb{E}[1_{\{\rho \geq \tau_1\}} v(X_\rho^{\tau_1, \xi_1, B^3, S^3}, Y_\rho^{\tau_1, \xi_1, B^3, S^3}) | \mathcal{G}_{\tau_1}] \\ &\geq 1_{\{\rho \geq \tau_1\}} v(\xi_1) = 1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) \end{aligned}$$

Taking \mathcal{G}_τ -conditional expectation, we get

$$(4.33) \quad \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \geq \mathbb{E}[1_{\{\rho \geq \tau_1\}} v^\eta(\xi_1) | \mathcal{G}_\tau].$$

Adding (4.31) and (4.33), we get

$$\mathbb{E}[v^\eta(X_\rho^{\tau, \xi, B, S}, Y_\rho^{\tau, \xi, B, S}) | \mathcal{G}_\tau] \geq v^\eta(\xi).$$

This completes the verification of (SB2) for v^η , and hence of the viscosity supersolution property of v_- . □

4.5 Comparison principle

A comparison principle can be established following the idea of [53]. The key is to show the existence of a strict subsolution which is then added to the penalty term when applying the technique of doubling of variables.

Lemma 4.5.1. *There exists a strict subsolution ℓ of (4.4) satisfying*

(1) $\ell \in C^2(\overline{\mathcal{S}})$ and $\ell < 0$;

(2) $\ell(x, y) \rightarrow -\infty$ as $\|(x, y)\| \rightarrow \infty$ in $\overline{\mathcal{S}}$.⁵

Proof. Let $h(z) := -\frac{(z-b+1)^p}{p}$ with $0 < p < 1$. We have $h < 0$, $h' < 0$ and $h'' > 0$ in $(b-1, \infty)$. Let $1 - \mu < k < \frac{1}{1-\lambda}$ and define $\ell(x, y) := h(x + ky)$. ℓ is well-defined since $x + ky \geq b$ for all $(x, y) \in \overline{\mathcal{S}}$. Condition (1) is trivially satisfied. To see condition (2) holds, observe that for each $a > b$, $\{(x, y) \in \overline{\mathcal{S}} : x + ky \leq a\}$ is a bounded subset of \mathbb{R}^2 . Therefore if $\|(x, y)\| \rightarrow \infty$ in $\overline{\mathcal{S}}$, then we must have $x + ky \rightarrow \infty$. It follows that $\ell(x, y) = h(x + ky) \rightarrow -\infty$. It remains to show ℓ is a strict subsolution of (4.4) under a suitable choice of p .

Let $(x, y) \in \mathcal{S}$. By our choice of k and that $h' < 0$, we readily obtain

$$-(1 - \mu)\ell_x + \ell_y = [-(1 - \mu) + k]h'(x + ky) < 0$$

and

$$\ell_x - (1 - \lambda)\ell_y = [1 - k(1 - \lambda)]h'(x + ky) < 0.$$

Let us now compute $\mathcal{L}\ell(x, y)$.

$$\begin{aligned} \mathcal{L}\ell(x, y) &= \beta\ell(x, y) - (rx - c)\ell_x(x, y) - \alpha y\ell_y(x, y) - \frac{1}{2}\sigma^2 y^2 \ell_{yy}(x, y) \\ &= \beta h(x + ky) - (rx - c + \alpha ky)h'(x + ky) - \frac{1}{2}\sigma^2 y^2 k^2 h''(x + ky). \end{aligned}$$

By definition of the solvency region \mathcal{S} , we have

$$x + (1 - \mu)y < \frac{c}{r} \quad \text{if } y > 0, \quad \text{and} \quad x + \frac{y}{1 - \lambda} < \frac{c}{r} \quad \text{if } y < 0,$$

which implies

$$rx - c + \alpha ky \leq \frac{r|y|}{1 - \lambda} + \alpha k|y| = \left(\frac{r}{1 - \lambda} + \alpha k \right) |y| := \theta|y|.$$

⁵The function ℓ is referred to as a Lyapunov function in [53].

Using $h'(x + ky) < 0$ and $h''(x + ky) > 0$, we deduce

$$\begin{aligned}
\mathcal{L}l(x, y) &\leq \beta h(x + ky) - \theta|y|h'(x + ky) - \frac{1}{2}\sigma^2 y^2 k^2 h''(x + ky) \\
&= -\frac{1}{2} \left[\sigma^2 y^2 k^2 h''(x + ky) + 2\theta|y|h'(x + ky) + \frac{\theta^2 (h'(x + ky))^2}{\sigma^2 k^2 h''(x + ky)} \right] \\
&\quad + \beta h(x + ky) + \frac{1}{2} \frac{\theta^2 (h'(x + ky))^2}{\sigma^2 k^2 h''(x + ky)} \\
&\leq \left(\beta + \frac{1}{2} \frac{\theta^2}{\sigma^2 k^2} \frac{(h')^2}{hh''}(x + ky) \right) h(x + ky) \\
&= \left(\beta - \frac{1}{2} \frac{\theta^2}{\sigma^2 k^2} \frac{p}{1-p} \right) h(x + ky).
\end{aligned}$$

Choose p small such that $\beta > \frac{1}{2} \frac{\theta^2}{\sigma^2 k^2} \frac{p}{1-p}$. We then have by negativity of h that

$$\mathcal{L}l(x, y) < 0. \quad \square$$

Proposition 4.5.2. *Let u, v be u.s.c. viscosity subsolution and l.s.c. viscosity supersolution of (4.4), respectively. Suppose u, v are bounded and $u \leq v$ on $\partial\mathcal{S}$, then $u \leq v$ in \mathcal{S} .*

Proof. Assume to the contrary that $\delta := u(x_0, y_0) - v(x_0, y_0) > 0$ for some $(x_0, y_0) \in \mathcal{S}$. Let ℓ be the strict classical subsolution given by Lemma 4.5.1. Let ϵ be a small positive constant satisfying $\delta + 2\epsilon\ell(x_0, y_0) > 0$. For each $\theta > 0$, define

$$\Phi_\theta(x, y, x', y') := u(x, y) - v(x', y') - \frac{\theta}{2}(|x - x'|^2 + |y - y'|^2) + \epsilon\ell(x, y) + \epsilon\ell(x', y').$$

Since $u(x, y) - v(x', y')$ is u.s.c. and bounded, and $\ell(x, y) \rightarrow -\infty$ as $\|(x, y)\| \rightarrow \infty$ in $\bar{\mathcal{S}}$, there exists $(x_\theta, y_\theta), (x'_\theta, y'_\theta)$ lying in a compact subset of $\bar{\mathcal{S}}$ such that

$$\sup_{(x, y), (x', y') \in \bar{\mathcal{S}}} \Phi_\theta(x, y, x', y') = \Phi_\theta(x_\theta, y_\theta, x'_\theta, y'_\theta).$$

Compactness allows us to extract a sequence $\theta_n \rightarrow \infty$ such that $(x_n, y_n, x'_n, y'_n) := (x_{\theta_n}, y_{\theta_n}, x'_{\theta_n}, y'_{\theta_n}) \rightarrow (\hat{x}, \hat{y}, \hat{x}', \hat{y}')$ as $n \rightarrow \infty$. Clearly, we have

$$(4.34) \quad \Phi_{\theta_n}(x_n, y_n, x'_n, y'_n) \geq \sup_{(x, y) \in \bar{\mathcal{S}}} \Phi_0(x, y, x, y) \geq \delta + 2\epsilon\ell(x_0, y_0) > 0.$$

It follows that

$$\begin{aligned} \frac{\theta_n}{2}(|x_n - x'_n|^2 + |y_n - y'_n|^2) &\leq u(x_n, y_n) - v(x'_n, y'_n) + \epsilon\ell(x_n, y_n) + \epsilon\ell(x'_n, y'_n) \\ &\quad - \sup_{(x,y) \in \bar{\mathcal{S}}} \Phi_0(x, y, x, y). \end{aligned}$$

Since the right hand side is bounded from above and $\theta_n \rightarrow \infty$, we must have $|x_n - x'_n|^2 + |y_n - y'_n|^2 \rightarrow 0$, hence $(\hat{x}, \hat{y}) = (\hat{x}', \hat{y}')$. This further implies by u.s.c. of $u - v$ that

$$0 \leq \limsup_n \frac{\theta_n}{2}(|x_n - x'_n|^2 + |y_n - y'_n|^2) \leq \Phi_0(\hat{x}, \hat{y}, \hat{x}, \hat{y}) - \sup_{(x,y) \in \bar{\mathcal{S}}} \Phi_0(x, y, x, y) \leq 0.$$

So we conclude

$$(4.35) \quad \lim_n \theta_n(|x_n - x'_n|^2 + |y_n - y'_n|^2) = 0,$$

and

$$(4.36) \quad \lim_n \Phi_{\theta_n}(x_n, y_n, x'_n, y'_n) = \Phi_0(\hat{x}, \hat{y}, \hat{x}, \hat{y}) = \sup_{(x,y) \in \bar{\mathcal{S}}} \Phi_0(x, y, x, y) > 0.$$

Now, since $u \leq v$ on $\partial\mathcal{S}$ and $\ell \leq 0$, we have $\Phi_0(x, y, x, y) \leq 0$ for $(x, y) \in \partial\mathcal{S}$. In view of (4.36), we have $(\hat{x}, \hat{y}) \in \mathcal{S}$. So $(x_n, y_n), (x'_n, y'_n) \in \mathcal{S}$ for n sufficiently large. By Crandall-Ishii's lemma, we can find matrices $A_n, B_n \in \mathbb{S}_2$ such that

$$(4.37) \quad (\theta_n(x_n - x'_n), \theta_n(y_n - y'_n), A_n) \in \bar{J}_S^{2,+}(u(x_n, y_n) + \epsilon\ell(x_n, y_n)),$$

$$(4.38) \quad (\theta_n(x_n - x'_n), \theta_n(y_n - y'_n), B_n) \in \bar{J}_S^{2,-}(v(x'_n, y'_n) - \epsilon\ell(x'_n, y'_n)),$$

and

$$\begin{pmatrix} A_n & 0 \\ 0 & -B_n \end{pmatrix} \leq 3\theta_n \begin{pmatrix} I & -I \\ -I & I \end{pmatrix}.$$

where $\bar{J}_S^{2,+}$ and $\bar{J}_S^{2,-}$ denote the closure of the second order superjet and subjet, respectively. By Lemma 4.2.7 of [55], we have

$$(4.39) \quad (y_n)^2 A_{n,22} - (y'_n)^2 B_{n,22} \leq 3\theta_n |y_n - y'_n|^2.$$

Since ℓ is a $C^2(\mathcal{S})$ functions, we can rewrite (4.37) and (4.38) as

$$(p_n, X_n) \in \bar{J}_S^{2,+} u(x_n, y_n), \quad (q_n, Y_n) \in \bar{J}_S^{2,-} v(x'_n, y'_n)$$

where

$$\begin{aligned} p_n &:= \theta_n(x_n - x'_n, y_n - y'_n) - \epsilon D\ell(x_n, y_n), \quad X_n := A_n - \epsilon D^2\ell(x_n, y_n), \\ q_n &:= \theta_n(x_n - x'_n, y_n - y'_n) + \epsilon D\ell(x'_n, y'_n), \quad Y_n := B_n + \epsilon D^2\ell(x'_n, y'_n). \end{aligned}$$

By the semijets definition of viscosity solution, we have

$$\begin{aligned} \max \left\{ \beta u(x_n, y_n) - (rx_n - c)p_{n,1} - \alpha y_n p_{n,2} - \frac{1}{2}\sigma^2 y_n^2 X_{n,22}, \right. \\ \left. - (1 - \mu)p_{n,1} + p_{n,2}, \quad p_{n,1} - (1 - \lambda)p_{n,2} \right\} \leq 0 \end{aligned}$$

and

$$\begin{aligned} \max \left\{ \beta v(x'_n, y'_n) - (rx'_n - c)q_{n,1} - \alpha y'_n q_{n,2} - \frac{1}{2}\sigma^2 (y'_n)^2 Y_{n,22}, \right. \\ \left. - (1 - \mu)q_{n,1} + q_{n,2}, \quad q_{n,1} - (1 - \lambda)q_{n,2} \right\} \geq 0. \end{aligned}$$

We consider three cases.

Case 1. $-(1 - \mu)q_{n,1} + q_{n,2} \geq 0$ for infinitely many n 's. In this case,

$$\begin{aligned} 0 &\geq -(1 - \mu)p_{n,1} + p_{n,2} - [-(1 - \mu)q_{n,1} + q_{n,2}] \\ &= -\epsilon[-(1 - \mu)\ell_x(x_n, y_n) + \ell_y(x_n, y_n)] - \epsilon[-(1 - \mu)\ell_x(x'_n, y'_n) + \ell_y(x'_n, y'_n)]. \end{aligned}$$

Letting $n \rightarrow \infty$ yields

$$0 \geq -2\epsilon[-(1 - \mu)\ell_x(\hat{x}, \hat{y}) + \ell_y(\hat{x}, \hat{y})],$$

or

$$-(1 - \mu)\ell_x(\hat{x}, \hat{y}) + \ell_y(\hat{x}, \hat{y}) \geq 0.$$

This is a contradiction to the strict subsolution property of ℓ in the sell region.

Case 2. $q_{n,1} - (1 - \lambda)q_{n,2} \geq 0$ for infinitely many n 's. Similar to case 1, this leads to $\ell_x(\hat{x}, \hat{y}) - (1 - \lambda)\ell_y(\hat{x}, \hat{y}) \geq 0$, contradicting the strict subsolution property of ℓ in the buy region.

Case 3. For n sufficiently large, $\beta v(x'_n, y'_n) - (rx'_n - c)q_{n,1} - \alpha y'_n q_{n,2} - \frac{1}{2}\sigma^2 (y'_n)^2 Y_{n,22} \geq 0$. In this case,

$$\begin{aligned} 0 &\leq \beta v(x'_n, y'_n) - (rx'_n - c)q_{n,1} - \alpha y'_n q_{n,2} - \frac{1}{2}\sigma^2 (y'_n)^2 Y_{n,22} \\ &\quad - \left[\beta u(x_n, y_n) - (rx_n - c)p_{n,1} - \alpha y_n p_{n,2} - \frac{1}{2}\sigma^2 y_n^2 X_{n,22} \right] \\ &= -\beta [u(x_n, y_n) - v(x'_n, y'_n)] + \epsilon(\mathcal{L}\ell - \beta\ell)(x_n, y_n) + \epsilon(\mathcal{L}\ell - \beta\ell)(x'_n, y'_n) \\ &\quad + r\theta_n(x_n - x'_n)^2 + \alpha\theta_n(y_n - y'_n)^2 + \frac{1}{2}\sigma^2 [y_n^2 A_{n,22} - (y'_n)^2 B_{n,22}] \\ &\leq -\beta [u(x_n, y_n) - v(x'_n, y'_n) + \epsilon\ell(x_n, y_n) + \epsilon\ell(x'_n, y'_n)] \\ &\quad + \left(r + \alpha + \frac{3}{2}\sigma^2 \right) \theta_n(|x_n - x'_n|^2 + |y_n - y'_n|^2) \\ &= -\beta\Phi_{\theta_n}(x_n, y_n, x'_n, y'_n) + \left(r + \alpha + \frac{3}{2}\sigma^2 - \frac{\beta}{2} \right) \theta_n(|x_n - x'_n|^2 + |y_n - y'_n|^2) \\ &\leq -\beta(\delta - 2\epsilon\ell(x_0, y_0)) + \left(r + \alpha + \frac{3}{2}\sigma^2 - \frac{\beta}{2} \right) \theta_n(|x_n - x'_n|^2 + |y_n - y'_n|^2). \end{aligned}$$

In the third step, we used the subsolution property of ℓ and (4.39). In the fourth step, we used the definition of Φ_θ . In the last step, we used (4.34). Letting $n \rightarrow \infty$ and using (4.35), we arrive at the contradiction $0 \leq -\beta(\delta - 2\epsilon\ell(x_0, y_0)) < 0$. The proof is complete. \square

4.5.1 Proof of Theorem 4.2.5.

Proof. By Remarks 4.3.3 and 4.4.3, we have $v_- \leq \psi \leq v_+$. By Propositions 4.3.6 and 4.4.7, we know v_+ is a viscosity subsolution and v_- is a viscosity supersolution of (4.4). Moreover, $v_+ \leq v_-$ on $\partial\mathcal{S}$. It is also clear that v_+ is u.s.c. and v_- is l.s.c.. Comparison principle (Proposition 4.5.2) then implies $v_+ \leq v_-$. Therefore, $v_+ = v_- = \psi$ is a continuous viscosity solution to the Dirichlet problem (4.4), (4.5). Uniqueness also follows from the comparison principle. \square

APPENDIX

APPENDIX A

A.1 Support of measures and random variables

Definition A.1.1. Let \mathcal{X} be a topological space. The support of a nonempty family of probability measures $\mathcal{R} \subseteq \mathfrak{P}(\mathcal{X})$, denoted by $\text{supp}(\mathcal{R})$, is the smallest closed \mathcal{R} -q.s. set in \mathcal{X} :

$$\text{supp}(\mathcal{R}) := \bigcap \{A \subset \mathcal{X} \text{ closed} : R(A) = 1 \forall R \in \mathcal{R}\}.$$

Equivalently, $\text{supp}(\mathcal{R})$ consists of the set of points $x \in \mathcal{X}$ such that for every open neighborhood N_x of x , there is a measure $R \in \mathcal{R}$ with $R(N_x) > 0$.

Definition A.1.2. Let $S : \Omega \rightarrow \mathbb{R}$ be a random variable, and $\mathcal{R} \subseteq \mathfrak{P}(\Omega)$ be a nonempty family of probability measures on Ω . The support of S under \mathcal{R} , denoted by $\text{supp}_{\mathcal{R}}S$, is defined as the support of the law of S under \mathcal{R} , i.e.

$$\text{supp}_{\mathcal{R}}S = \text{supp}\{P \circ S^{-1} : P \in \mathcal{R}\}.$$

One can show that $\text{supp}_{\mathcal{R}}S$ is the smallest closed set $A \subseteq \mathbb{R}$ such that $P(S \in A) = 1 \forall P \in \mathcal{R}$. Equivalently, a point y belongs to $\text{supp}_{\mathcal{R}}S$ if and only if every open ball around y has positive measure under some member of $\{P \circ S^{-1} : P \in \mathcal{R}\}$.

The next lemma is used in the construction of the modified market in Sections 2.3.3.

Lemma A.1.3. *Let $S : \Omega \rightarrow \mathbb{R}$ be continuous and $\mathcal{R} \subseteq \mathfrak{P}(\Omega)$ be a family of probability measures. Then*

$$\text{supp}_{\mathcal{R}} S(\cdot) = \overline{S(\text{supp}(\mathcal{R}))}.$$

Proof. Let $y \in \overline{S(\text{supp}(\mathcal{R}))}$. Then $\forall r > 0$, $\exists \omega_r \in \text{supp}(\mathcal{R})$ such that $S(\omega_r) \in B_r(y)$. Since S is continuous, $S^{-1}(B_r(y))$ is an open neighborhood of ω_r . Since $\omega_r \in \text{supp}(\mathcal{R})$, $\exists P_r \in \mathcal{R}$ such that $P_r \circ S^{-1}(B_r(y)) > 0$. We therefore have $y \in \text{supp}_{\mathcal{P}} S(\cdot)$. The other inclusion does not require S to be continuous. Suppose $y \notin \overline{S(\text{supp}(\mathcal{R}))}$. Then $\exists r > 0$ such that $B_r(y) \cap \overline{S(\text{supp}(\mathcal{R}))} = \emptyset$, which implies $S^{-1}(B_r(y)) \cap \text{supp}(\mathcal{R}) = \emptyset$. It follows that $\forall P \in \mathcal{R}$, we must have $P \circ S^{-1}(B_r(y)) = 0$, otherwise $P(\text{supp}(\mathcal{R}))$ would be strictly less than 1, contradicting the definition of $\text{supp}(\mathcal{R})$. So the neighborhood $B_r(y)$ is \mathcal{R} -polar, meaning $y \notin \text{supp}_{\mathcal{R}} S(\cdot)$. \square

A.2 Continuous and measurable set-valued maps

Definition A.2.1. Let $\Phi : \Omega \rightarrow \mathcal{X}$ be a set-valued map between topological spaces.

We say Φ is

- (i) upper hemicontinuous at ω if for every open neighborhood U of $\Phi(\omega)$, the upper inverse $\{\omega \in \Omega : \Phi(\omega) \subseteq U\}$ includes an open neighborhood of ω . Φ is upper hemicontinuous (on Ω) if it is upper hemicontinuous at every $\omega \in \Omega$.
- (ii) lower hemicontinuous at ω if for every open set U which meets $\Phi(\omega)$, the lower inverse $\{\omega \in \Omega : \Phi(\omega) \cap U \neq \emptyset\}$ includes an open neighborhood of ω . Φ is lower hemicontinuous (on Ω) if it is lower hemicontinuous at every $\omega \in \Omega$.
- (iii) continuous if it is both upper and lower hemicontinuous.

Definition A.2.2. Let (Ω, Σ) be a measurable space and \mathcal{X} be a topological space.

A set-valued map $\Phi : (\Omega, \Sigma) \rightarrow \mathcal{X}$ is measurable (resp. weakly measurable) if

$\Phi^{-1}(B) := \{\omega \in \Omega : \Phi(\omega) \cap B \neq \emptyset\}$ ¹ is measurable for each closed (resp. open) subset B of \mathcal{X} .

When \mathcal{X} is σ -compact separable metrizable space \mathcal{X} (e.g. \mathbb{R}^n) and Φ is closed-valued. Then measurable and weakly measurable are equivalent [43, Theorem 3.2(ii)]. We will use the words *universally measurable* when $\Sigma = \mathcal{F}$ and *Borel measurable* when $\Sigma = \mathcal{B}$.

In the proof of the FTAP, the construction of a dual element, under the local-to-global philosophy, relies heavily on measurable selection theorems. We include a version of the famous Jankov-von Neumann selection theorem here for the reader's convenience (see e.g. [76, Theorem 5.5.2] or [22, Proposition 7.49]).

Theorem A.2.3 (Jankov-von Neumann theorem). *Let X, Y be Polish spaces and $A \subseteq X \times Y$ be an analytic set. Then there exists a universally measurable function $\phi : \text{proj}_X(A) \rightarrow Y$ such that $\text{graph}(\phi) \subseteq A$.*

A.3 A useful lemma on equivalent change of measure

The next lemma is taken from [34, Theorem VII.57] (see also [25, Lemma 3.2]). It says integrability of a sequence of random variables can be obtained by an equivalent change of measure.

Lemma A.3.1. *Let $P \in \mathfrak{P}(\Omega)$ and f_n be a sequence of (P -a.s. finite) random variables. There exists probability measure $R \sim P$ with bounded density with respect to P , such that all f_n are R -integrable.*

¹This notion of inverse is sometimes called *lower inverse* or *weak inverse*.

A.4 Proof of Lemma 2.3.17

Proof. Let $\tau := \min\{t : -\xi_t \in K_t \setminus K_t^0\}$. We are done if $\tau = \infty$ \mathcal{P} -q.s.. Suppose $\exists P \in \mathcal{P}$ such that $P(\tau < \infty) > 0$. Let $[\underline{S}', \overline{S}'] \subseteq ri[\underline{S}, \overline{S}]$ satisfy $\text{NA}(\mathcal{P})$ and let K'_t denote the solvency cone for the $[\underline{S}', \overline{S}']$ -market. We have $K_t \subseteq K'_t$. Let $\epsilon_t = \epsilon_t(\omega)$ be the largest (nonnegative) number such that $\xi_t(\omega) + \epsilon_t(\omega)\mathbf{1} \in -K'_t(\omega)$. We have $\epsilon_t > 0$ on $\{\tau = t\}$. It can also be shown that $\epsilon_t \in \mathcal{F}_t$. Define

$$\eta_t := \begin{cases} \xi_t, & t \neq \tau, \\ \xi_t + \epsilon_t \mathbf{1}, & t = \tau. \end{cases}$$

Then $\eta_t \in L^0(-K'_t; \mathcal{F}_t)$. Since $\sum_{t=0}^T \eta_t = \sum_{t=0}^T 1_{\{t=\tau\}} \epsilon_t \mathbf{1} \geq 0$ \mathcal{P} -q.s., $\text{NA}(\mathcal{P})$ for the $[\underline{S}', \overline{S}']$ -market implies $\sum_{t=0}^T 1_{\{t=\tau\}} \epsilon_t \mathbf{1} = 0$ \mathcal{P} -q.s., in particular, P -a.s., contradicting the fact that $P(\tau < \infty) > 0$ and $\epsilon_t(\omega) > 0$ on $\{\tau = t\}$. \square

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer. A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. ArXiv Preprint, version 2, to appear in *Mathematical Finance*.
- [2] A. D. Alexandroff. Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it. *Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser.*, 6:3–35, 1939.
- [3] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis : a hitchhiker's guide*. Springer-Verlag, Berlin Heidelberg, 3rd edition, 2006.
- [4] O. Alvarez, J.-M. Lasry, and P.-L. Lions. Convex viscosity solutions and state constraints. *J. Math. Pures Appl. (9)*, 76(3):265–288, 1997.
- [5] N. Bäuerle and E. Bayraktar. A note on applications of stochastic ordering to control problems in insurance and finance. *Stochastics*, 86(2):330–340, 2014.
- [6] E. Bayraktar, A. Cosso, and H. Pham. Robust feedback switching control: dynamic programming and viscosity solutions. 2014. ArXiv Preprint, version 1.
- [7] E. Bayraktar, X. Hu, and V. R. Young. Minimizing the probability of lifetime ruin under stochastic volatility. *Insurance Math. Econom.*, 49(2):194–206, 2011.
- [8] E. Bayraktar and J. Li. Stochastic perron for stochastic target games. 2014. ArXiv Preprint, version 3. To appear in the *Annals of Applied Probability*.
- [9] E. Bayraktar and M. Sîrbu. Stochastic Perron's method and verification without smoothness using viscosity comparison: the linear case. *Proc. Amer. Math. Soc.*, 140(10):3645–3654, 2012.
- [10] E. Bayraktar and M. Sîrbu. Stochastic Perron's method for Hamilton-Jacobi-Bellman equations. *SIAM J. Control Optim.*, 51(6):4274–4294, 2013.
- [11] E. Bayraktar and M. Sîrbu. Stochastic Perron's method and verification without smoothness using viscosity comparison: obstacle problems and dynkin games. *Proc. Amer. Math. Soc.*, 142(4):1399–1412, 2014.

- [12] E. Bayraktar and S. Yao. A weak dynamic programming principle for zero-sum stochastic differential games with unbounded controls. *SIAM J. Control Optim.*, 51(3):2036–2080, 2013.
- [13] E. Bayraktar and V. R. Young. Correspondence between lifetime minimum wealth and utility of consumption. *Finance Stoch.*, 11(2):213–236, 2007.
- [14] E. Bayraktar and V. R. Young. Minimizing the probability of lifetime ruin under borrowing constraints. *Insurance Math. Econom.*, 41(1):196–221, 2007.
- [15] E. Bayraktar and V. R. Young. Minimizing the probability of ruin when consumption is ratcheted. *N. Am. Actuar. J.*, 12(4):428–442, 2008.
- [16] E. Bayraktar and V. R. Young. Proving regularity of the minimal probability of ruin via a game of stopping and control. *Finance Stoch.*, 15(4):785–818, 2011.
- [17] E. Bayraktar and Y. Zhang. Fundamental theorem of asset pricing under transaction costs and model uncertainty. 2013. ArXiv Preprint, version 2.
- [18] E. Bayraktar and Y. Zhang. Minimizing the Probability of Lifetime Ruin Under Ambiguity Aversion. *SIAM J. Control Optim.*, 53(1):58–90, 2015.
- [19] E. Bayraktar and Y. Zhang. Stochastic Perron’s Method for the Probability of Lifetime Ruin Problem Under Transaction Costs. *SIAM J. Control Optim.*, 53(1):91–113, 2015.
- [20] E. Bayraktar, Y. Zhang, and Z. Zhou. A note on the fundamental theorem of asset pricing under model uncertainty. *Risks*, 2(4):425–433, 2014.
- [21] A. Ben-Tal, D. Bertsimas, and D. B. Brown. A soft robust model for optimization under ambiguity. *Oper. Res.*, 58(4, part 2):1220–1234, 2010.
- [22] D. P. Bertsekas and Shreve S. E. *Stochastic Optimal Control*. Academic Press, New York, 1978.
- [23] G. Bordigoni, A. Matoussi, and M. Schweizer. A stochastic control approach to a robust utility maximization problem. In *Stochastic analysis and applications*, volume 2 of *Abel Symp.*, pages 125–151. Springer, Berlin, 2007.
- [24] B. Bouchard and M. Nutz. Consistent Price Systems under Model Uncertainty. ArXiv Preprint, version 1.
- [25] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.*, 25(2):823–859, 2015.
- [26] S. Browne. Survival and growth with a liability: optimal portfolio strategies in continuous time. *Math. Oper. Res.*, 22(2):468–493, 1997.
- [27] A. Cartea, R. Donnelly, and S. Jaimungal. Robust market making. 2013. Available at SSRN: <http://ssrn.com/abstract=2310645>.

- [28] L. Cousot. Conditions on option prices for absence of arbitrage and exact calibration. *Journal of Banking & Finance*, 31(11):3377 – 3397, 2007.
- [29] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc., New Ser.*, 27(1):1–67, 1992.
- [30] R. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics Stochastics Rep.*, 29(2):185–201, 1990.
- [31] M. H. A. Davis and D. G. Hobson. The range of traded option prices. *Math. Finance*, 17(1):1–14, 2007.
- [32] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4):676–713, 1990.
- [33] M. H. A. Davis, J. Obłój, and V. Raval. Arbitrage bounds for prices of weighted variance swaps. *Math. Finance*, 24(4):821–854, 2014.
- [34] C. Dellacherie and P. A. Meyer. *Probabilities and Potential B*. North Holland, Amsterdam, 1982.
- [35] Y. Dolinsky and H. M. Soner. Robust hedging with proportional transaction costs. *Finance Stoch.*, 18(2):327–347, 2014.
- [36] D. Duffie, W. H. Fleming, H. M. Soner, and T. Zariphopoulou. Hedging in incomplete markets with HARA utility. *J. Econom. Dynam. Control*, 21(4-5):753–782, 1997.
- [37] W. H. Fleming and D. Hernández-Hernández. On the value of stochastic differential games. *Commun. Stoch. Anal.*, 5(2):341–351, 2011.
- [38] W. H. Fleming and P. E. Souganidis. On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana Univ. Math. J.*, 38(2):293–314, 1989.
- [39] P. Grigoriev. On low dimensional case in the fundamental asset pricing theorem with transaction costs. *Statist. Decisions*, 23(1):33–48, 2005.
- [40] L. P. Hansen and T. J. Sargent. *Robustness*. Princeton University Press, 2007.
- [41] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. *J. Econom. Theory*, 128(1):45–90, 2006.
- [42] D. Hernández-Hernández and A. Schied. A control approach to robust utility maximization with logarithmic utility and time-consistent penalties. *Stochastic Process. Appl.*, 117(8):980–1000, 2007.
- [43] C. J. Himmelberg. Measurable relations. *Fund. Math.*, 87:53–72, 1975.

- [44] Y. Hu and M. Schweizer. Some new BSDE results for an infinite-horizon stochastic control problem. In *Advanced mathematical methods for finance*, pages 367–395. Springer, Heidelberg, 2011.
- [45] C. Huang and H. Pagès. Optimal consumption and portfolio policies with an infinite horizon: existence and convergence. *Ann. Appl. Probab.*, 2(1):36–64, 1992.
- [46] H. Ishii. Perron’s method for Hamilton-Jacobi equations. *Duke Math. J.*, 55(2):369–384, 1987.
- [47] S. Jacka. Avoiding the origin: a finite-fuel stochastic control problem. *Ann. Appl. Probab.*, 12(4):1378–1389, 2002.
- [48] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 2nd edition, 2002.
- [49] S. Jaimungal. Irreversible investments and ambiguity aversion. 2011. Available at SSRN: <http://ssrn.com/abstract=1961786>.
- [50] S. Jaimungal and G. Sigloch. Incorporating risk and ambiguity aversion into a hybrid model of default. *Math. Finance*, 22(1):57–81, 2012.
- [51] K. Janeček and M. Sîrbu. Optimal investment with high-watermark performance fee. *SIAM J. Control Optim.*, 50(2):790–819, 2012.
- [52] Y. Kabanov. In discrete time a local martingale is a martingale under an equivalent probability measure. *Finance Stoch.*, 12(3):293–297, 2008.
- [53] Y. Kabanov and C. Klüppelberg. A geometric approach to portfolio optimization in models with transaction costs. *Finance Stoch.*, 8(2):207–227, 2004.
- [54] Y. Kabanov, M. Rásonyi, and C. Stricker. No-arbitrage criteria for financial markets with efficient friction. *Finance Stoch.*, 6(3):371–382, 2002.
- [55] Y. Kabanov and M. Safarian. *Markets with Transaction Costs, Mathematical Theory*. Springer-Verlag Berlin Heidelberg, 2009.
- [56] Y. Kabanov and C. Stricker. The Harrison-Pliska arbitrage pricing theorem under transaction costs. *J. Math. Econom.*, 35(2):185–196, 2001.
- [57] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 2nd edition, 1991.
- [58] N. V. Krylov. *Nonlinear Elliptic and Parabolic Equations of the Second Order*. D. Reidel Publishing Company, Dordrecht, Holland, 1987.
- [59] A. E. B. Lim and J. G. Shanthikumar. Relative entropy, exponential utility, and robust dynamic pricing. *Oper. Res.*, 55(2):198–214, 2007.

- [60] P. Maenhout. Robust portfolio rules and asset pricing. *The Review of Financial Studies*, 17(4):951–983, 2004.
- [61] S. Mataramvura and B. Øksendal. Risk minimizing portfolios and HJBI equations for stochastic differential games. *Stochastics*, 80(4):317–337, 2008.
- [62] M. A. Milevsky and C. Robinson. Self-annuitization and ruin in retirement. *N. Am. Actuar. J.*, 4(4):112–129, 2000. With discussion.
- [63] V. C. Pestien and W. D. Sudderth. Continuous-time red and black: how to control a diffusion to a goal. *Math. Oper. Res.*, 10(4):599–611, 1985.
- [64] L. C. G. Rogers. *Optimal Investment*. Springer Berlin Heidelberg, 2013.
- [65] D. B. Rokhlin. The martingale selection problem in the case of finite discrete time. *Teor. Veroyatn. Primen.*, 50(3):480–500, 2005.
- [66] D. B. Rokhlin. A constructive no-arbitrage criterion with transaction costs in the case of finite discrete time. *Teor. Veroyatn. Primen.*, 52(1):41–59, 2007.
- [67] D. B. Rokhlin. Martingale selection problem and asset pricing in finite discrete time. *Electron. Comm. Probab.*, 12:1–8, 2007.
- [68] D. B. Rokhlin. Verification by stochastic Perron’s method in stochastic exit time control problems. *J. Math. Anal. Appl.*, 419(1):433–446, 2014.
- [69] D. B. Rokhlin and G. Mironenko. Regular finite fuel stochastic control problems with exit time. 2015. ArXiv Preprint, version 1.
- [70] Dmitry B. Rokhlin. Stochastic Perron’s method for optimal control problems with state constraints. *Electron. Commun. Probab.*, 19:no. 73, 15, 2014.
- [71] W. Schachermayer. The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. *Math. Finance*, 14(1):19–48, 2004.
- [72] A. Schied. Optimal investments for risk- and ambiguity-averse preferences: a duality approach. *Finance Stoch.*, 11(1):107–129, 2007.
- [73] S. E. Shreve and H. M. Soner. Optimal investment and consumption with transaction costs. *Ann. Appl. Probab.*, 4(3):609–692, 1994.
- [74] M. Sîrbu. Stochastic Perron’s method and elementary strategies for zero-sum differential games. *SIAM J. Control Optim.*, 52(3):1693–1711, 2014.
- [75] M. Smaga. *Utility-based proof for the existence of strictly consistent price processes under proportional transaction costs*. PhD thesis, Technischen Universität Kaiserslautern, 2012.

- [76] S. M. Srivastava. *A course on Borel sets*, volume 180 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.
- [77] D. W. Stroock. *Lectures on Stochastic Analysis: Diffusion Theory*. Cambridge University Press, 1987.
- [78] A. E. Taylor. L'Hospital's rule. *Amer. Math. Monthly*, 59:20–24, 1952.
- [79] A. Tourin and T. Zariphopoulou. Portfolio selection with transaction costs. In *Seminar on Stochastic Analysis, Random Fields and Applications (Ascona, 1993)*, volume 36 of *Progr. Probab.*, pages 385–391. Birkhäuser, Basel, 1995.
- [80] N. Touzi. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*. Springer New York Heidelberg Dordrecht London, 2012.
- [81] R. Uppal and T. Wang. Model misspecification and underdiversification. *The Journal of Finance*, 58(6):2465–2486, 2003.
- [82] V. R. Young. Optimal investment strategy to minimize the probability of lifetime ruin. *N. Am. Actuar. J.*, 8(4):105–126, 2004.
- [83] T. Zariphopoulou. Consumption-investment models with constraints. *SIAM J. Control Optim.*, 32(1):59–85, 1994.