University of Trento, Department of Mathematics

Financial risk sources and optimal strategies in jump-diffusion frameworks

PhD thesis in Mathematics

Luca Prezioso

Assessment Committee

Luca Di Persio	University of Verona	Advisor
Viktor Bezborodov	Wroclaw University	Referee
Stefano Bonaccorsi	University of Trento	Examiner
Lucian Maticiuc	Alexandru Ioan Cuza University	Referee and Examiner
Adrian Zălinescu	Alexandru Ioan Cuza University	Referee and Examiner

March 25, 2020

To Anna

Acknowledgments

Firstly, I would like to thank my beloved Anna, indeed without her this achievement would probably not have been possible and certainly would have been immensely more difficult. Day after day she encouraged me continuously and was always there for me, ready to patiently advise me. She made me feel her closeness, even in the years spent abroad, and even more in the difficult times. I realized that even if you work hard, if the person you love is far away, the results will not be so good as you hope. With your loved one close to you, the outcomes will be much better.

I would like to express my gratitude to my advisor and mentor Prof. Luca di Persio for his support and never ending patience and time dedicated to me. During these three years of my PhD he has been my role model in the academic field: an example of dedication, knowledge and humility, from whom I certainly still have much to learn. His guidance helped me on several occasions and he opened many, many doors for me, from academic collaborations to useful professional contacts. It is to him that I owe the chance I had to speak in important conferences in Europe, USA and Asia, and the opportunity I had to join a great team at the forefront of probabilistic studies in Paris for to years. I could not have had a better advisor for my PhD studies.

I would like to thank also the rest of my thesis committee: Prof. Adrian Zălinescu, Prof. Lucian Maticiuc, Prof. Viktor Bezborodov and Prof. Stefano Bonaccorsi, for the time they took to review the thesis and for their discerning comments and suggestions.

During the three years of my PhD I got the chance to work for a great team within Allianz Global Investors, its members: Dhikra, Michaela, Tilman and Tobias are people that I would like to thank. Especially their leader Dr. Kai Wallbaum is one of the greatest people I had the opportunity to meet. He was always a great support to me and he made me realize what a great environment the professional world is, helping me to move from an academic world perspective to a more real world one.

Among the collaborators of some of the projects included in this thesis I would particularly like to thank Dr. Francesco Cordoni, whose intellectual generosity I was most impressed by.

Last but not least, I have to thank my family and my friends for their contributions to this important achievement. Especially my mother, who is the person who most encouraged me to pursue this path from the beginning. It was she who taught me to always strive for excellence and made me become the person I am now. My father also played an important part, making me often notice the reality behind the beautiful facades. I would like to thank also my brothers Matteo and Davide, and my grandparents Rosemarie and Hans. A particular thought goes to my grandfather, the polar star of this journey of mine and who unfortunately I could not have beside me.

The period spent in Paris was the most important for the finalisation of my thesis. These two years will always remain a good memory, thanks to the person that most of all I discovered to be a true friend. In addition to being a dedicated researcher, Fabio helped me a lot to integrate into a different culture, but even more helped me to feel at home in Paris.

I would also like to thank Anna's parents: Luciana and Tonino, for their closeness and generosity, and for having always been of great support.

Among the other people who played a role in this thesis, I thank: Davide, Sara and Alessandro, and nonna Elvira; Yilun from Penn State University; Assaf, Côme and Guillame from Paris Diderot University; Nicolò and Francesco from the University of Trento; and Nicolò, Florian, Marco and Jacopo.

Contents

Thesis Overview

The thesis considers different sources of risk that affect financial business in general, and in particular the banks' value. It is divided as follows: Part I treats an optimal dividend problem with investment opportunities, taking into consideration a source of strategic risk; Part II and III concern the wide concept of credit risk from two different perspectives, namely considering the systemic risk feature and the regulatory risk point of view; part IV and V concern the financial risk in general, and in particular the volatility risk. Notably the last part is devoted to the implementation of exact formulas for a strategy designed to reduce the volatility risk of the investment strategy.

Let me briefly explain the main results and characteristics of each part, taking the opportunity to list the co-authors. More introductory details will be introduced at the beginning of each part.

Part I: Strategic Risk and Dividend Control Problem under Financing Constraints *This part results from the collaboration with the Laboratoire de Probabilités, Statistique et Modélisation at the university of Paris Diderot, in particular with Assoc. Prof. Scotti and Prof. Ly Vath.*

In the first part, the effect of market frictions on the decision process of the financial entities is considered. It concerns the problem of determining an optimal control of the dividend under debt constraints and investment opportunities in an economy with business cycles. It is assumed that the company is to be allowed to accept or reject investment opportunities arriving at random times with random sizes, by changing its outstanding indebtedness, which would impact its capital structure and risk profile. This work mainly focuses on the strategic risk faced by the companies; and, in particular, it focuses on the manager's problem of setting appropriate priorities to deploy the limited resources available. This component is taken into account by introducing frictions in the capital structure modification process.

The problem is formulated as a bi-dimensional singular control problem under regime switching in presence of jumps. An explicit condition is obtained in order to ensure that the value function is finite. A viscosity solution approach is used to get qualitative descriptions of the solution.

Furthermore, in the second chapter of this part it is shown that the bi-dimensional problem could be reduced, by writing the value function and the processes under the debt value. The reduced value function is characterized and used to perform complete comparative statics.

Part II: A lending scheme for systemic risk with probabilistic constraints of failure *This part results from the collaboration with Assoc. Prof. di Persio and Dr. Cordoni of the probability team at the university of Verona.*

Financial institutions cannot possibly carry enough capital to withstand counterparty failures or systemic risk. In such situations, the central bank or the government becomes effectively the risk manager of last resort or, in extreme cases, the lender of last resort. If, on the one hand, the health of the whole financial system depends on government intervention, on the other hand, guaranteeing a high probability of salvage may result in increasing the moral hazard of the banks in the financial network.

In this part we derive a closed form solution for an optimal control problem related to interbank lending schemes, subject to terminal probability constraints on the failure of banks which are interconnected through a financial network. The derived solution applies to real bank networks by obtaining a general solution when the aforementioned probability constraints are assumed for all the banks. We also present a direct method to compute the systemic relevance parameter for each bank within the network.

Part III: The Default Risk Charge approach to regulatory risk measurement processes *This part results from the collaboration with Assoc. Prof. di Persio and Dr. Bonollo.*

This part considers the Default Risk Charge measure as an effective alternative to the Incremental Risk Charge one, proposing its implementation by a quasi exhaustive-heuristic algorithm to determine the minimum capital requested to a bank facing the market risk associated to portfolios based on assets emitted by several financial agents. While most of the banks use the Monte Carlo simulation approach and the empirical quantile to estimate this risk measure, we provide new computational approaches, exhaustive or heuristic, currently becoming feasible, because of both new regulation and the high speed - low cost technology available nowadays. Concrete algorithms and numerical examples are provided to illustrate the effectiveness of the proposed techniques.

Part IV: Portfolio optimization for Hawkes-Merton models with transaction costs *This part results from the collaboration with Assoc. Prof. Scotti and Assoc. Prof. Sgarra.*

A financial market is deemed to be composed of two investment opportunities: a *risk-free asset*, also called money market account or Government bond or simply bond, which evolves at a risk-free interest rate, and a *risky asset*, also referred to as stock or share, subject to a significant degree of volatility and a jump component driven by a Hawkes process. We investigate the Optimal Consumption-Investment problem when proportional transaction costs are taken into account. The solution of the problem, which is stochastic, is proven to be the deterministic solution of a second-order integro-differential equation in the viscosity sense, and a detailed qualitative analysis of the solution obtained is provided.

We show that the global effects of the combination of self-exciting jumps and proportional transaction costs is to increase the investors' fear whenever they perceive a hint of financial crisis, in other words we take into account the emotional states of irrational market behavior, which is inconsistent with the efficient market hypothesis.

Part V: Closed formula for options linked to target volatility strategies *This part results from the collaboration with Assoc. Prof. di Persio and Dr. Wallbaum, Head of Global Asset-Life Solution (Allianz GI).*

In recent years we have seen a new class of structured products emerging, which made use of options linked to dynamic asset allocation strategies. One of the most chosen approaches is the so-called "target volatility mechanism", which shifts between risky and riskless asset, aiming to control the volatility of the overall portfolio. This strategy aims to reduce the volatility risk in portfolios of derivative instruments.

There are a series of articles looking into options linked to the target volatility mechanism, but this study is the first, which tries to develop closed-end formulas for VolTarget options. In a Black-Scholes environment we develop closed-end formulas for option prices and some key hedging parameters, when the underlying is following a target volatility mechanism.

Part I

Strategic Risk and Dividend Control Problem under Financing Constraints

Financial peace isn't the acquisition of stuff. It's learning to live on less than you make, so you can give money back and have money to invest. You can't win until you do this.

Dave Ramsey (1960-)

Introduction

Firm managers have often to solve a dilemma about the cash flows. The shareholder's are waiting for dividends but the same money could be kept to invest and make future cash flows. The celebrated paper of Modigliani and Miller [92] addresses the problem and shows that the value of a firm is unaffected by how that firm is financed, that is cash reserves are irrelevant. However, this capital structure irrelevance principle is crucially based on frictionless hypothesis. In particular, managers could cover funding for future projects without costs and capital supply is totally elastic. Empirical studies indicate that the credit suppliers, instead firm managers, have the upper hand in financing decisions, see for instance Graham and Harvey [62]. Moreover, other studies highlight the precautionary role of the cash holdings by pointing that cash ratios have more than doubled during the last decades, see for instance Bates et al. [14].

There is a vast and increasing literature on firm's dividend and investment policy in a stochastic environment, see for instance [35, 47, 52, 67, 71, 116] and Dixit and Pindick [49] for a survey. The optimal strategy for the firm is generally characterized by means of stopping times defined by the time when the cash process reaches an endogenous threshold obtained as free boundary of a variational problem. Probably the first attempt to study dividend policies without external constraint is due to Jeanblanc and Shiryaev [71] where singular stochastic control theory is used in a Brownian framework. Chevalier et al. [35] propose a model for dividend and investment policy of a firm under debt constraints.

The firm, we consider in these two chapters, has a cash reserve following a geometric Brownian motion as in the Merton model. As in Chevalier et al. [35], we assume that the firm carries a debt obligation on its balance sheet. In this part of the thesis we as well will assume that firm assets is cash equivalent as in the literature mainstream. We allow firm's manager to pay dividends and to leverage firm capital structure by debt raising. However, the firm cannot increase debt straightaway but needs to search for credit suppliers, that is we introduce liquidity risk in the manager's optimization problem. We model the related matching times via a Poisson process. Mathematically, we formulated the problem as a combined singular and optimal control problem. The related Hamilton Jacobi Bellman system is a variational partial integro-differential equation.

Our main contribution is to give an analytical solution to the firm optimal policy in presence of external financing frictions, in other words the hypothesis of perfect elastic credit supply is relaxed, but in a weak sense, since the inelasticity will not be money but only time consuming. For instance, Graham and Harvey [62] found little evidence that managers are concerned about transaction costs.

We enrich our analysis including a business cycle into the evolution of cash flows. The uncertainty conveyed by the business cycle to the cash flow has an evident impact on the cash holding and, combined with credit frictions, could explain the increasing size of cash ratio of firms. The credit crisis of 2008 gives an illustration, Campello et al. [32] found that the inability to borrow externally caused many firms to bypass attractive investment opportunities. Therefore, the analysis of this part will focus on the strategic risk faced by the company's decision maker, that is the source of loss that may arise from an unsuccessful business plan, e.g. dividend payments, wrong timing of investment activities, inadequate resource allocation, or from a failure to respond well to changes in the business environment. Another analysis about the impact of credit constraints on the behavior of real firms is proposed by Duchin et al [51] finding that corporate investment declines significantly following the onset of the crisis. Consistent with a causal effect of a supply shock, the decline is greater for firms that have low cash reserves.

The firm, we consider, in this part of the thesis, must decide when it is optimal to pay dividends and debt modifications with business cycle uncertainty and credit supply frictions. We consider in particular search frictions in a similar way as Hugonnier et al. [67, 68], see also Villeneuve and Warin [116]. We first show that the value function is finite if and only if the discount factor is larger than an explicit threshold different to the average value of the growth rates. We characterize the optimal policy and the value function in term of the unique viscosity solution to the associated system of quasi-variational integro partial differential inequalities. Then, assuming that the debt interest rate does not depend on the debt level, we show that the value function can be written using the firm debt as unit of account reducing the dimension to a one-dimensional problem. The auxiliary value function can be characterized via a system of variational inequality. We study numerically the sensitivity of the auxiliary value function.

Our main result, in financial point of view, is to reconcile the seminal model of Jeanblanc and Shiryaev [71] with Modigliani and Miller [92] cash irrelevance. As a matter of fact, our comparative statics highlight that the optimal dividend threshold decreases as the frequency of external financing opportunities arrival. When the intensity is zero, our model could be seen as a Merton version of Jeanblanc and Shiryaev one. When the intensity grows to infinity, the optimal threshold falls to the level such that the equity is negligible with respect to the debt. This result could in particular explain the increase of cash ratios pointed by Bates et al. [14] linking it with the credit crunch.

The remainder of this part is organized as follows. We define the model describing the decision variables in the first chapter. Section 2 is devoted to the characterization of the solution of the problem in terms of the unique viscosity solution to the associated HJB system and to obtain some qualitative results about this solution. In Section 4 an approximation scheme of the HJB is provided through a Picard's type procedure. Chapter 2 presents the dimension reduction result showing that the value function could be deduced using an auxiliary value function, solution of a one-dimensional HJB, and discusses empirical predictions of the markets and the comparative statics.

Chapter 1

The model

We consider an economy with risk neutral agents that discount the cash flow at a fixed rate $\rho > 0$. The uncertainty is described by a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, with right-continuous filtration. An admissible control strategy will be $\alpha = \big(Z_t, \{\pi_k\}_{k\in\mathbb{N}}\big)$, where Z is a non-decreasing càdlàg process representing the dividend policy, and $\{\pi_k\}$ is $\{0,1\}$-valued, where its values 0 and 1 correspond respectively to the rejection or the acceptance of the k-th investment or withdraw event and it is measurable with respect to the filtration at the arrival time. Moreover we consider admissible strategies avoiding the asset value going below the debt value.

1.1 Model formulation

We consider a firm which has assets denoted by $\{X_t\}_{t\in\mathbb{R}^+}$. This assets are assumed cash equivalent. We assume that the firm is conditioned by a business cycle, that is the drift of cash flows depends on an external random source. We consider that the business cycle is driven by a continuous time homogeneous Markov chain $\{M_t\}_{t\in\mathbb{R}^+}$ with state space $\mathcal{S} = \{0,1\}$: the expansion period in which the economy faces growth and increased production and the contraction period characterized by slowed growth. When the contraction reaches the trough, the economy starts the recovery and a new cycle begins. Historically, form the economical perspective on this topic and considering short business cycles of about 3-5 years it is worth noting the Kitching business cycle, see e.g., [73], or, see e.g., [105], for an argument concerning general length periods of the economic waves and more updated on the present financial situation. The related transition matrix $\Lambda = (\Lambda_{h,k})_{h,k=0,1}$, where $\Lambda_{i,1-i}$ is the intensity of transition form state i to state $1-i$. Namely the economic phase at time t, denoted by M_t, is the current state of the Markov chain: assumed to be continuous time homogeneous, and generated by its transition matrix $\Lambda = (\Lambda_{h,k})_{h,k=0,1}$, where $\Lambda_{h,k}$ is the constant intensity of transition form state h to state k.

$$\lim_{\Delta t \to 0} \frac{\mathbb{P}(M_{t+\Delta t} = k \mid M_t = h, \{M_s\}_{0 \le s < t})}{\Delta t} = \Lambda_{k,h}\,, \tag{1.1.1}$$

while for $h \in \{0,1\}$, $\Lambda_{h,h} = -\Lambda_{h,1-h}$; this implies that the lifetime of each state is exponentially distributed with parameter $\Lambda_{h,h}$. Denote by $j \in \{0,1\}$ the starting value for M, and by τ_n the time of the n-th switch in the business cycle, hence $M_{\tau_n} = \frac{1+(-1)^{n+j-1}}{2}$ for each $n \in \mathbb{N}_+$, i.e. for odd n, then $M_{\tau_n} = j$, otherwise $M_{\tau_n} = 1-j$.

The firm assets generates a continuous stream of cash flow satisfying a Black Scholes model

$$\mathrm{d}X_t = b_{M_t}\, X_t \,\mathrm{d}t + \sigma\, X_t \,\mathrm{d}W_t\,.$$

The Asset-Liability Model is referred to the balance sheet theory, where in particular, as defined by the Accounting Standard Board, by asset we mean a resource controlled by the enterprise as a result of past events and from which future economic benefits are expected to flow in the enterprise, and on the other hand the liabilities are the future sacrifices of economic benefits that the enterprise is presently obliged to make to other entities as a result of past transactions or past events. In other words, poor business decisions made by the manager of the company can lead to losses, both realized losses and losses of future possible profits. That is the company has to take into account the *strategic risk*, a leading factor in modern risk management.

Since assets are assumed cash equivalent, we consider that current liabilities are negligible. Long liabilities value, hereinafter debt, is denoted by $\{Y_t\}_{t\in\mathbb{R}^+}$ and then shareholders' equity worth $\{X_t - Y_t\}_{t\in\mathbb{R}^+}$. In order to simplify the model, we assume that debt is not going to expire, instead the firm has to pay interests continuously. Unlike assets, credit markets exhibit frictions, which implies that the firm manager needs to look for credit providers. We assume that this search is time consuming but there is no issuance or cancellation costs for debt changes. The firm meets a credit supplier at the random arrival times of a Poisson process N with intensity rate $\lambda > 0$. The sequence of random times will be denoted by $\{\theta_i\}_{i\in\mathbb{N}}$. However, credit suppliers have finite financial resources, we suppose that the firm manager can choose to accept or refuse the proposal of the credit supplier. We consider that the size of the proposed debt is proportional to the actual level of the debt. We justify this choice in a reputation acquisition point of view, see Diamond [48]. A possible interpretation of our assumption is that the previous debt could be used by credit suppliers as a measure of the reliability of the firm. Moreover, we assume that credit suppliers can also propose to withdraw a part of the debt previously provided and we furthermore assume that firm's manager can accept or refuse bearing in mind that the credit suppliers and the manager herself have signed an agreement for a perennial debt and then a new agreement is needed. We also suppose that the credit supplier has finite financial resources then she proposes a size, denoted by $\{\zeta_i\}_{i\in\mathbb{N}}$ taking value on $(-1,\infty)\backslash\{0\}$, where a positive (resp. negative) ζ constitutes a incremental (resp. withdraw) offer. The firm's acceptance/refusal is described by the sequence $\{\pi_i\}_{i\in\mathbb{N}}$. With these definition the evolution of the debt reads

$$Y_t = y \prod_{k=1}^{\infty} (1 + \zeta_k \, \pi_k \, \mathbb{1}_{\theta_k \le t}) \, , \quad \text{for } t \ge 0 \, , \tag{1.1.2}$$

assuming the initial value $Y_0 = y > 0$.

The firm is liquidate when the equity value reaches zero and then the manager refunds the debt holders in advance. We then denote by T^α, or simply T when no ambiguity occurs, the bankruptcy time

$$T^\alpha = \inf\big\{ t \ge 0 : \; X_t^{(j,x,y,\alpha)} \le Y_t^{(j,x,y,\alpha)} \big\} \, , \tag{1.1.3}$$

i.e. we assume that the firm faces bankruptcy whenever its asset value goes beyond its liabilities, the so called *balance sheet insolvency*.

We assume that the interest rate r is a non-decreasing bounded function of the debt level y. The repayment of the debt in advance is contrary to the initial agreement of perennial loan and then it affects the reliability of the manager. Then we add a penalty for the firm manager at bankruptcy proportional to the level of the debt, in a similar way of Eaton and Gersovitz [54]. This penalty P is non negative and it could be included the case of a loss given default rate. We will see the implication of no penalty in the last section of the next chapter assuming a constant interest rate, the two hypotheses can be seen as a limit of no credit risk.

The firm's asset dynamics, denoted by a process X, is then governed by the following stochastic differential equation (SDE):

$$\mathrm{d}X_t = (b_{M_t} \, X_t - r(Y_t) \, Y_t) \, \mathrm{d}t + \sigma \, X_t \, \mathrm{d}W_t - \mathrm{d}Z_t + \mathrm{d}Y_t \, , \quad \forall t \ge 0, \tag{1.1.4}$$

with $X_0 = x$. By convention we set the ordering $b_0 < b_1$, where b_0 represents the drift of the firm's asset value in the contraction period and b_1 its drift in the expansion period.

The maximization problem has the following value function

$$v_j(x,y) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,x,y)}\left[\int_0^{T^-} e^{-\rho t}\,\mathrm{d}Z_t - P\,Y_T\,e^{-\rho T}\right], \quad (x,y)\in D,\ M_0 = j\,, \tag{1.1.5}$$

where the subscript of v corresponds to the starting economic phase, ρ is the discount rate, the domain $D := \{(x,y)|0 < y < x\}$, and $\mathcal{A}$ is the set of admissible control strategies

$$\mathcal{A} := \{\alpha = (Z_t, \{\pi_k\}_{k\in\mathbb{N}^+})\ :\ 0 < Y_t^\alpha < X_t^\alpha,\ \forall\, t > 0\}\,,$$

i.e. an admissible strategy is adapted to the filtration and it avoids the enterprise to pay dividends that exceed the debt value.

1.2 Dynamic Programming Principle and related HJB equation

Let us introduce the auxiliary process $S_t^{(j,x,y)}$ solution to

$$\mathrm{d}S_t = b_j\,S_t\,\mathrm{d}t + \sigma\,S_t\,\mathrm{d}W_t - r(y)\,y\,\mathrm{d}t, \quad \text{for } t\in\mathbb{R}^+, \tag{1.2.1}$$

with $S_0 = x$ and $j\in\{0,1\}$, where the first two terms of the right-hand side represent the asset value generation, and the last term the initial liability cost.

We may now assume the following Dynamic Programming Principle: for every stopping time η and every $(x,y)\in D$, we have

$$v_j(x,y) = \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,x,y)}\left[\int_0^{T^-\wedge\eta} e^{-\rho t}\,\mathrm{d}Z_t + e^{-\rho\,(T^-\wedge\eta)} v_{M_{T^-\wedge\eta}}(X_{T^-\wedge\eta}, Y_{T^-\wedge\eta})\right], \tag{1.2.2}$$

by using the dynamic programming principle we obtain the system of variational inequalities satisfied by the value functions:

$$\begin{cases} \min\Big\{\rho\,v_j(x,y) - \mathcal{L}_j v_j(x,y) - \mathcal{J} v_j(x,y) - \mathcal{G}_j v_j(x,y)\,;\ \frac{\partial v_j(x,y)}{\partial x} - 1\Big\} = 0\,, \\ v_j(y,y) = -P\,y\,, \end{cases} \tag{1.2.3}$$

for $(x,y)\in D$ and $j\in\{0,1\}$, where we have the following operators

$$\begin{aligned} \mathcal{L}_j v_j(x,y) &= \big(b_j\,x - r(y)\,y\big)\frac{\partial v_j}{\partial x}(x,y) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 v_j}{\partial x^2}(x,y)\,, \\ \mathcal{G}_j v_j(x,y) &= \Lambda_{j,1-j}\big(v_{1-j}(x,y) - v_j(x,y)\big)\,, \end{aligned}$$

$$\mathcal{J} v_j(x,y) \;=\; \lambda \int_{-1}^{\infty} \big(v_j(x+y\,\zeta, y\,(\zeta+1)) - v_j(x,y)\big)_+ \,\nu(\mathrm{d}\zeta)\,.$$

1.3 Properties of value function and viscosity characterization

In this section, we deal with the characterization of the value function via viscosity theory. We start by showing some basic properties of value function and a necessary condition (1.3.9) for the discount rate to avoid divergence in the value function. Condition (1.3.9) is not standard since it depends on both the growth rates b_j and the regime switching rates $\Lambda_{j,1-j}$ in a non linear way. Then we prove a comparison principle guaranteeing a sufficient condition for being a supersolution to the original HJB equation (1.2.3). Moreover we show that the value functions satisfy a linear growth condition under the same condition (1.3.9) that is then also sufficient and that they are continuous on their whole domain. Finally, the main result of this section is Theorem 1.3.6, where we prove that the value functions are indeed the unique viscosity solution to HJB (1.2.3).

1.3.1 Basic properties

Lemma 1.3.1 *The value function is*

- *increasing with respect to x;*
- *decreasing with respect to y and*
- *$v_0(x,y) \le v_1(x,y)$ for all $(x,y) \in D$.*

Proof. Assume $x_1 < x_2$ and fix $0 < \epsilon < x_2 - x_1$, consider the strategy consisting in distributing at initial time $x_2 - x_1$ as dividends and then follow an ϵ-optimal strategy for the initial condition (x_1, y). Thanks to the DPP we have $v_i(x_2, y) \ge x_2 - x_1 + v_i(x_1, y) - \epsilon > v_i(x_1, y)$.

Assume $y_1 < y_2$ and consider ϵ small enough, let α_ϵ be an ϵ-optimal strategy for the initial condition (x, y_2), we will denote by T_2^α the related bankruptcy time. If $T_2^\alpha = 0$ then we just consider the liquidation strategy also for the initial condition (x, y_1) and, supposing $\epsilon < y_2 - y_1$, we obtain easily $v_i(x, y_1) \ge y_2 - y_1 + v_i(x, y_2) - \epsilon > v_i(x, y_2)$. Otherwise, we consider the initial condition (x, y_1) and a strategy $\overline{\alpha}_\epsilon$ that consists in following the strategy α_ϵ adding the continuous rate of dividend $r(y_2)\,y_2 - r(y_1)\,y_1$

up to T_2^α. It is easy to remark that this strategy is admissible and that the related bankruptcy time T_1^α is larger than T_2^α, thanks to the DPP we have

$$v_i(x, y_1) \geq [r(y_2)\, y_2 - r(y_1)\, y_1]\;\mathbb{E}\,[T_2^\alpha] + v_i(x, y_2) - \epsilon$$

since the expectation of T_2^α is strictly positive and that the previous inequality is true for any positive ϵ we obtain $v_i(x, y_1) > v_i(x, y_2)$.

Finally fix the same initial condition (x, y) with two regimes and consider ϵ small enough, let α_ϵ be an ϵ-optimal strategy for the regime 0, we will denote by T_0^α the related bankruptcy time. If $T_0^\alpha = 0$ the result is evident, we then consider $T_0^\alpha > 0$. We introduce $\overline{\tau} := T_0^\alpha \wedge \tau_1^{(0)} \wedge \tau_1^{(1)}$, where $\tau_1^{(i)}$ denotes the first transition time between regime i to $1 - i$. We consider now the regime 1 and a strategy $\overline{\alpha}_\epsilon$ that is to follow the strategy α_ϵ adding the continuous rate of dividend $(b_1 - b_0)X_t$ up to $\overline{\tau}$. It is easy to check that the strategy is admissible and that the related bankruptcy time T_1^α is not smaller than $\overline{\tau}$ and, a fortiori, than $\overline{\tau} \wedge \theta_1$. Thanks to the DPP we have

$$\begin{aligned} v_1(x, y) &\geq \mathbb{E}\left[\int_0^{\overline{\tau}\wedge\theta_1} (b_1 - b_0)\, X_t\, \mathrm{d}t\right] + v_0(x, y_2) - \epsilon \\ &\geq (b_1 - b_0)\, y\, \mathbb{E}\,[\overline{\tau} \wedge \theta_1] + v_0(x, y_2) - \epsilon \end{aligned}$$

since the expectation of T_2^α is strictly positive and that the previous inequality is true for any positive ϵ, we obtain $v_1(x, y) \wedge v_0(x, y)$. □

1.3.2 Necessary condition for non-divergent value functions

Lemma 1.3.2 *Let $\rho < \max\{b_1 - \Lambda_{1,0}, b_0 - \Lambda_{0,1}\}$ then $v_j(x, y)$ diverges to infinity for all $(x, y) \in D$.*

Proof. We first consider the case $b_1 - \Lambda_{1,0} > b_0 - \Lambda_{0,1}$ and $\rho < b_1 - \Lambda_{1,0}$. We consider as starting regime 1, let τ_1^1 the first switching time from regime 1 to 0. We introduce a restriction $\overline{\mathcal{A}}$ on the class of admissible strategies $\mathcal{A}$, such that for all $\theta_i \leq \tau_1^1$, $\pi_i = 0$ and $Z_{\tau_1^1} = X_{\tau_1^1}^{j,x,y} - y$. That means that the class $\overline{\mathcal{A}}$ is such that the manager refuses all investment/disinvestment opportunities and she returns all the cash to investors at time τ_1^1 liquidating the firm. We can define the value function $u_1(x, y)$ associated to this problem via the definition (1.1.5) restricted to $\overline{\mathcal{A}}$. The associated HJB is

$$\min\Big\{(\rho+\Lambda_{1,0})\,u_1(x,y)-(b_1x-r(y)y)\frac{\partial u_1}{\partial x}(x,y)-\frac{1}{2}\sigma^2x^2\frac{\partial^2 u_1}{\partial x^2}(x,y)-\Lambda_{1,0}(x-y-Py),\\ \frac{\partial u_1}{\partial x}-1\Big\}=0$$

with the condition $u_1(y,y)=-Py$. This HJB is the same of an optimal dividend payment over infinite horizon with discount rate $\rho+\Lambda_{1,0}$ smaller than the growth rate b by hypothesis. It is well-known that the associated value function $u_1(x,y)$ equals infinity for all $x>y$. Since $\overline{\mathcal{A}}\subset\mathcal{A}$ we have that $v_1(x,y)\geq u_1(x,y)$ and then $v_1(x,y)$ equals infinity for all $x>y$.

We now consider the regime 0, applying the dynamic programming principle (7.3.1) choosing $\eta=\tau_1^0$ that is the first regime switching time from regime 0 to 1, we have

$$v_0(x,y)=\sup_{\alpha\in\mathcal{A}}\mathbb{E}^{(0,x,y)}\Bigg[\int_0^{T^-\wedge\tau_1^0}e^{-\rho\,t}\,\mathrm{d}Z_t+e^{-\rho\,T^-}v_0(X_{T^-},Y_{T^-})\mathbb{1}_{T^-\leq\tau_1^0}\\ +e^{-\rho\,\tau_1^0}v_1(X_{\tau_1^0},Y_{\tau_1^0})\mathbb{1}_{T^->\tau_1^0}\Bigg].$$

Since the last term is infinite by the previous analysis and that the indicator is not always zero, we have $v_0(x,y)$ diverges to infinity for all $x>y$.

A symmetric argument gives the result if $b_1-\Lambda_{1,0}<b_0-\Lambda_{0,1}$ and $\rho<b_0-\Lambda_{0,1}$. $\square$

Proposition 1.3.1 *Let*

$$\widetilde{b}:=\left(\frac{b_0}{\Lambda_{0,1}}+\frac{b_1}{\Lambda_{1,0}}\right)\left(\frac{1}{\Lambda_{0,1}}+\frac{1}{\Lambda_{1,0}}\right)^{-1}=\frac{\Lambda_{1,0}\,b_0+\Lambda_{0,1}\,b_1}{\Lambda_{1,0}+\Lambda_{0,1}}$$

and assume $\rho<\widetilde{b}$, then $v_j(x,y)$ diverges to infinity $(x,y)\in D$, $j\in\{0,1\}$.

Proof. Let $\rho<\widetilde{b}$. We now fix $\epsilon>0$ such that $\rho<\widetilde{b}-\epsilon$. Consider the following sequence of admissible strategies with

$$Z_t^{(k)}=\begin{cases}0 & \text{if } t<t_k\\ X_{t_k^-}-y & \text{if } t=t_k\end{cases},\qquad \pi_i=0,\quad\forall i$$

$k\in\mathbb{N}^+$, where $\{t_k\}_{k\geq1}$ is strictly non-decreasing and tends to infinity when k tends to infinity, i.e. the strategy $\alpha_k = (Z_t^{(k)}, 0)$ consists in rejecting every opportunity and waiting up to t_k to pay everything in dividends, hence for each strategy α_k the bankruptcy time is $T = t_k$. The corresponding asset process $X^{(j,x,y,\alpha_k)}$, denoted by $X^{(k)}$, satisfies the SDE

$$\mathrm{d}X_t^{(k)} = X_t^{(k)}\big(b_{M_t}\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\big) - r(y)\,y\,\mathrm{d}t\,, \quad \text{for } 0 < t < t_k\,, \tag{1.3.1}$$

$X_0^{(k)} = x$. Let $C_t := \int_0^t b_{M_s}\,\mathrm{d}s + \sigma\,W_t$, and $H_t := x - r(y)\,y\,t$, using [103, Thm. 52 Sec. 9 Ch. V], we may rewrite (1.3.5) as

$$\mathrm{d}X_t^{(k)} = X_t^{(k)}\,\mathrm{d}C_t + \mathrm{d}H_t\,, \quad \text{for } 0 < t < t_k\,,$$

and, defining the Doléans-Dade exponential $\mathcal{E}(C)_t = \exp\{\int_0^t b_{M_s}\,\mathrm{d}s + \sigma\,W_t - \frac{1}{2}\sigma^2\,t\}$, by [103] we have the following expression for $X_t^{(k)}$ for $0 < t < t_k$

$$\begin{aligned} X_t^{(k)} &= \mathcal{E}(C)_t\left(x - \int_0^t \frac{r(y)\,y}{\mathcal{E}(C)_s}\mathrm{d}s\right)\\ &= x\,U_t\,\exp\left(\int_0^t b_{M_s}\,\mathrm{d}s\right) - r(y)\,y\int_0^t \frac{U_t}{U_s}\exp\left(\int_s^t b_{M_u}\,\mathrm{d}u\right)\mathrm{d}s\,, \end{aligned}$$

with $U_t = \exp(\sigma\,W_t - \frac{1}{2}\sigma^2\,t)$. Therefore, taking the expectation, by Fubini theorem and the independence of $\exp(\int_0^t b_{M_s}\,\mathrm{d}s)$ and U_t, we have

$$\mathbb{E}\left[X_t^{(k)}\right] = x\,\mathbb{E}\left[\exp\left(\int_0^t b_{M_s}\,\mathrm{d}s\right)\right] - r(y)\,y\int_0^t \mathbb{E}\left[\exp\left(\int_s^t b_{M_u}\,\mathrm{d}u\right)\right]\mathrm{d}s\,. \tag{1.3.2}$$

By denoting $\Phi(s,t) := \mathbb{E}[\exp(\int_s^t b_{M_u}\mathrm{d}u)]$ for every $0 \le s \le t$, and discounting the expectation in (1.3.6), we get the following expression

$$\mathbb{E}\left[e^{-\rho\,t}\,X_t^{(k)}\right] = e^{-\rho\,t}\,\Phi(0,t)\left(x - r(y)\,y\int_0^t \frac{\Phi(s,t)}{\Phi(0,t)}\mathrm{d}s\right), \quad \text{for } 0 < t < t_k\,, \tag{1.3.3}$$

We first compute

$$\begin{aligned} e^{-\rho\,t}\,\Phi(0,t) &= e^{-\rho\,t}\,\mathbb{E}\left[\exp\left(\int_0^t b_{M_u}\mathrm{d}u\right)\right]\\ &\ge e^{-\rho\,t}\exp\left(\mathbb{E}\left[\int_0^t b_{M_u}\mathrm{d}u\right]\right)\\ &\ge e^{-(\rho-\widetilde{b})\,t}\exp\left(\mathbb{E}\left[\int_0^t \left(b_{M_u} - \widetilde{b}\right)\,\mathrm{d}u\right]\right) \end{aligned}$$

$$\geq \exp\left(\left(\epsilon + \frac{1}{t}\,\mathbb{E}\left[\int_0^t \left(b_{M_u} - \widetilde{b}\right)\,\mathrm{d}u\right]\right) t\right) ,$$

since $\frac{1}{t}\mathbb{E}\left[\int_0^t \left(b_{M_u} - \widetilde{b}\right)\,\mathrm{d}u\right]$ tends to zero when t tends to infinity, we have that $e^{-\rho\,t}\,\Phi(0,t)$ tends to infinity when t tends to infinity.

We now consider the second term in (1.3.3). Let $\underline{x}(y) := r(y)\,y\lim_{t\to\infty}\int_0^t \frac{\Phi(s,t)}{\Phi(0,t)}\,\mathrm{d}s$, which is bounded from above

$$\underline{x}(y) \leq \lim_{t\to\infty}\int_0^t \mathbb{E}[\exp(-b_0\,s)]\,\mathrm{d}s = r(y)\,y\lim_{t\to\infty}\frac{1}{b_0}\left(1 - e^{-b_0\,t}\right) = \frac{r(y)\,y}{b_0},$$

by simply noticing that $\Phi(0,t) \geq \Phi(s,t)\,e^{b_0\,s}$.

So we distinguish two cases:

- $x \geq \underline{x}(y) + \epsilon$, then the admissible strategy α_k gives a lower bound to the value function which tends to infinity when $k \to \infty$. Then, we have shown that the value function $v_j(x,y)$ is infinite for all $x \geq \underline{x}(y) + \epsilon$ and $j = 0, 1$.

- $y < x < \underline{x}(y) + \epsilon$, and introduce $\overline{\eta} := \inf_t\{X_t \geq \underline{x}(y) + \epsilon\}$, using the dynamic programming principle we may obtain that

$$\begin{aligned} v_j(x,y) &\geq \mathbb{E}\left[e^{-\rho\,\overline{\eta}}\,v_{M_\eta}(X_\eta, y)\,\mathbb{1}_{\overline{\eta}<T} - P\,e^{-\rho\,T}\,\mathbb{1}_{\overline{\eta}\geq T}\right] \\ &\geq \mathbb{E}\left[e^{-\rho\,\overline{\eta}}\,\mathbb{1}_{\overline{\eta}<T}\right] v_0(\underline{x}(y) + \epsilon, y) - P\,. \end{aligned}$$

  Since it is well known that the expected value $\mathbb{E}\left[e^{-\rho\overline{\eta}}\mathbb{1}_{\overline{\eta}<T}\right]$ is strictly positive, we have that the value function $v_j(x,y)$ equals infinity thanks to the first case.

□

We can now state a more general result about the constraint on the discount rate.

Proposition 1.3.2 *Let* $\rho > \max\{b_1 - \Lambda_{1,0}, b_0 - \Lambda_{0,1}\}$ *and assume that* ρ *is such that*

$$(\rho + \Lambda_{1,0} - b_1)\,(\rho + \Lambda_{0,1} - b_0) < \Lambda_{1,0}\Lambda_{0,1}\,, \tag{1.3.4}$$

then $v_j(x,y)$ *equals infinity for every* $(x,y) \in D$.

Proof. Consider the following sequence α_k of admissible strategies with $\{\pi_i = 0\}_{i\in\mathbb{N}}$ and

$$Z_t^{(k)} = \begin{cases} 0 & \text{if } t < \tau_{2k} \\ X_{\tau_{2k}^-} - y & \text{if } t = \tau_{2k} \end{cases}$$

Hence for each strategy α_k the bankruptcy time is $T = \tau_{2k}$. The corresponding asset process $X^{(j,x,y,\alpha_k)}$, denoted by $X^{(k)}$, satisfies the SDE

$$\mathrm{d}X_t^{(k)} = X_t^{(k)}\big(b_{M_t}\,\mathrm{d}t + \sigma\,\mathrm{d}W_t\big) - r(y)\,y\,\mathrm{d}t\,, \quad X_0^{(k)} = x\,. \tag{1.3.5}$$

Let $C_t := \int_0^t b_{M_s}\,\mathrm{d}s + \sigma\,W_t$, and $H_t := x - r(y)\,y\,t$, using Protter[103, Thm. 52 Sec. 9 Ch. V], we may rewrite (1.3.5) as

$$\mathrm{d}X_t^{(k)} = X_t^{(k)}\,\mathrm{d}C_t + \mathrm{d}H_t\,, \quad \text{for } 0 < t < t_k\,,$$

and, defining the Doléans-Dade exponential $\mathcal{E}(C)_t = \exp\{\int_0^t b_{M_s}\,\mathrm{d}s + \sigma\,W_t - \frac{1}{2}\sigma^2\,t\}$. By the method of "variation of constants", see for instance Section V.9 in [103], we have the following expression for $X_t^{(k)}$ for $0 < t < \tau_{2k}$

$$X_t^{(k)} = \mathcal{E}(C)_t\left(x - \int_0^t \frac{r(y)\,y}{\mathcal{E}(C)_s}\mathrm{d}s\right) = x\,R_t\,\exp\left(\int_0^t b_{M_s}\,\mathrm{d}s\right) - r\,y\int_0^t \frac{R_t}{R_s}\exp\left(\int_s^t b_{M_u}\,\mathrm{d}u\right)\mathrm{d}s\,,$$

with $R_t = \exp(\sigma\,W_t - \frac{1}{2}\sigma^2\,t)$. Therefore, taking the expectation, by Fubini theorem and the independence of $\exp(\int_0^t b_{M_s}\,\mathrm{d}s)$ and R_t, we have

$$\begin{aligned}\mathbb{E}\left[e^{-\rho\,t}\,X_t^{(k)}\right] &= x\,\mathbb{E}\left[\exp\left(-\int_0^t (\rho - b_{M_s})\,\mathrm{d}s\right)\right] - r(y)\,y\,e^{-\rho\,t}\int_0^t \mathbb{E}\left[\exp\left(\int_s^t b_{M_u}\,\mathrm{d}u\right)\right]\mathrm{d}s \\ &= \mathbb{E}\left[\exp\left(-\int_0^t (\rho - b_{M_s})\,\mathrm{d}s\right)\right]\left\{x - r(y)\,y\int_0^t \mathbb{E}\left[\exp\left(-\int_0^s b_{M_u}\,\mathrm{d}u\right)\right]\mathrm{d}s\right\}.\end{aligned} \tag{1.3.6}$$

We focus on $t = \tau_{2k}$ and we consider the first term in (1.3.6), we have

$$\mathbb{E}\left[\exp\left(-\int_0^{\tau_{2k}} (\rho - b_{M_s})\,\mathrm{d}s\right)\right] \tag{1.3.7}$$

$$\begin{aligned}&= \mathbb{E}\left[\exp\left\{-\sum_{j=1}^k\left(\int_{\tau_{2j-2}}^{\tau_{2j-1}} (\rho - b_{M_s})\,\mathrm{d}s + \int_{\tau_{2j-1}}^{\tau_{2j}} (\rho - b_{M_s})\,\mathrm{d}s\right)\right\}\right] \\ &= \prod_{j=1}^k \mathbb{E}\left[e^{-(\tau_{2j-1}-\tau_{2j-2})(\rho-b_0)-(\tau_{2j}-\tau_{2j-1})(\rho-b_1)}\right] \\ &= \left(\frac{\Lambda_{0,1}}{\rho + \Lambda_{0,1} - b_0}\;\frac{\Lambda_{1,0}}{\rho + \Lambda_{1,0} - b_1}\right)^k.\end{aligned} \tag{1.3.8}$$

Thanks to (1.3.4), the first term in (1.3.6) tends to infinity when k goes to infinity.

We now consider the second term in (1.3.6). Let 's define

$$\underline{x}(y) := r(y)\, y \lim_{k\to\infty} \int_0^{\tau_{2k}} \mathbb{E}\left[\exp\left(-\int_0^s b_{M_u}\, \mathrm{d}u\right)\right] \mathrm{d}s\,,$$

which is bounded from above

$$\underline{x}(y) \le \lim_{k\to\infty} \int_0^{\tau_{2k}} \mathbb{E}[\exp(-b_0\, s)]\, \mathrm{d}s = r(y)\, y \lim_{k\to\infty} \frac{1}{b_0}\left(1 - e^{-b_0\,\tau_{2k}}\right) = \frac{r(y)\, y}{b_0}\,.$$

So we distinguish two cases:

- $x \ge \underline{x}(y) + \epsilon$, then the admissible strategy α_k gives a lower bound to the value function which diverges to infinity when $k \to \infty$. Then, we have shown that the value function $v_j(x, y)$ is infinite for all $x \ge \underline{x}(y) + \epsilon$ and $j = 0, 1$.

- $y < x < \underline{x}(y) + \epsilon$, and introduce $\overline{\eta} := \inf\{t \,|\, X_t \ge \underline{x}(y) + \epsilon\}$, using the dynamic programming principle we may obtain that

$$\begin{aligned} v_j(x,y) &\ge \mathbb{E}\left[e^{-\rho\,\overline{\eta}}\, v_{M_\eta}(X_\eta, y)\, \mathbb{1}_{\overline{\eta}<T} - P\, y\, e^{-\rho T}\, \mathbb{1}_{\overline{\eta}\ge T}\right] \\ &\ge \mathbb{E}\left[e^{-\rho\,\overline{\eta}}\, \mathbb{1}_{\overline{\eta}<T}\right]\, v_0\left(\underline{x}(y) + \epsilon, y\right) - P\, y\,. \end{aligned}$$

  Since it is well known that the expected value $\mathbb{E}\left[e^{-\rho\overline{\eta}}\mathbb{1}_{\overline{\eta}<T}\right]$ is strictly positive, we have that the value function $v_j(x, y)$ equals infinity thanks to the first case.

□

To avoid the value function to be infinite, from now on we will assume that the discount rate satisfies $\rho > \max\{b_1 - \Lambda_{1,0}, b_0 - \Lambda_{0,1}\}$ and

$$(\rho + \Lambda_{1,0} - b_1)\,(\rho + \Lambda_{0,1} - b_0) \ge \Lambda_{1,0}\,\Lambda_{0,1}\,. \tag{1.3.9}$$

We remark that this condition is more restrictive to the usual one defined by Proposition 1.3.2. Moreover, as shown in Proposition 1.3.1, the discount rate ρ has to be larger than the long run mean value of the drift $\widetilde{b} := \left(\frac{b_0}{\Lambda_{0,1}} + \frac{b_1}{\Lambda_{1,0}}\right)\left(\frac{1}{\Lambda_{0,1}} + \frac{1}{\Lambda_{1,0}}\right)^{-1} = \frac{\Lambda_{1,0}b_0 + \Lambda_{0,1}b_1}{\Lambda_{1,0}+\Lambda_{0,1}}$. A direct computation shows that condition (1.3.9) is more restrictive than $\rho > \widetilde{b}$. Condition $\rho > \widetilde{b}$ could be seen as a linear constraint with respect to the growth rate, whereas condition (1.3.9) could be read as a geometric, and then non-linear, condition between growth and discounting in the spirit of relation (1.3.8). Our next objective is to show that this condition is not only necessary but also sufficient to have that the value function is finite. For that, we need some intermediate results.

1.3.3 Comparison principle and supersolution

Proposition 1.3.3 *Let* $\{\varphi_j\}_{j\in\{0,1\}} \in C^{2,0}(D)$ *be such that* $\varphi_j(y^+,y) := \lim_{x\downarrow y}\varphi_j(x,y) \geq -Py$, $\varphi_j(x,y) = -Py$ *for* $x<y$, *and*

$$\min\Big\{\rho\,\varphi_j - \mathcal{L}_j\varphi_j - \mathcal{J}\varphi_j - \mathcal{G}_j\varphi_j\,;\ \frac{\partial\varphi_j(x,y)}{\partial x} - 1\Big\} \geq 0\,, \tag{1.3.10}$$

then φ, *classical supersolution for the HJB equation* (1.2.3), *dominates the value function, that is* $\varphi_j \geq v_j$ *for all* $j\in\{0,1\}$ *and all* $(x,y)\in D$.

Proof. Fix (x,y) and $j\in\{0,1\}$, and consider the admissible strategy $\alpha = \big(Z,\{\pi_k\}_{k\in\mathbb{N}}\big)\in\mathcal{A}$. Let $m>0$, $k\in\mathbb{N}_0$, we consider the time t such that

$$T\wedge\theta_k \ \leq\ t\ <\ \widetilde{t}_{m,k} := T\wedge\theta_{k+1}\wedge\kappa_m\,,$$

where $\kappa_m := \inf\{t\geq T\wedge\theta_k\ :\ X_t^{(j,x,y,\alpha)}\geq m \text{ or } X_t^{(j,x,y,\alpha)}\leq Y_t^{(j,x,y,\alpha)}+\frac{1}{m}\}$. Hence define $\boldsymbol{\varphi} := [\varphi_0,\varphi_1]^T$, and apply Itô formula to $e^{-\rho t}\,\boldsymbol{\varphi}(X_t,Y_t)$ we have

$$\begin{aligned}
e^{-\rho\,\widetilde{t}_{m,k}}\,\boldsymbol{\varphi}(X_{\widetilde{t}^-_{m,k}},Y_{\widetilde{t}^-_{m,k}}) =& e^{-\rho\,(T\wedge\theta_k)}\,\boldsymbol{\varphi}(X_{T\wedge\theta_k},Y_{T\wedge\theta_k})\\
&+\int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\,(-\rho\,\boldsymbol{\varphi}+\mathrm{L}\boldsymbol{\varphi}+\mathcal{J}\boldsymbol{\varphi})(X_t,Y_t)\,\mathrm{d}t\\
&+\int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho t}\sigma X_t\frac{\partial\boldsymbol{\varphi}}{\partial x}(X_t,Y_t)\mathrm{d}W_t - \int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho t}\frac{\partial\boldsymbol{\varphi}}{\partial x}(X_t,Y_t)\mathrm{d}Z^c_t\\
&+\sum_{T\wedge\theta_k<t<\widetilde{t}_{m,k}} e^{-\rho\,t}\,\big(\boldsymbol{\varphi}(X_t,Y_t)-\boldsymbol{\varphi}(X_{t^-},Y_{t^-})\big)\\
&+\int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\,\mathrm{J}_2\,\boldsymbol{\varphi}(X_t,Y_t)\,\mathrm{d}M_t - \int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\,\mathcal{J}\boldsymbol{\varphi}(X_t,Y_t)\,\mathrm{d}t\,,
\end{aligned} \tag{1.3.11}$$

where $\mathrm{L} := \mathrm{diag}(\mathcal{L}_0,\mathcal{L}_1)$, $\mathrm{J}_2 := [0,1;1,0]$ is the 2-dimensional exchanging matrix, the fourth term results from the dividends continuously distributed, i.e. Z^c is the continuous part of Z. The fifth term arises from possible jumps in X_t due to the dividend distribution, i.e. ΔZ. We remark the fact that the martingale associated to the fore-last term in equation (1.3.11) reads as

$$\int \mathrm{J}_2\,\boldsymbol{\varphi}(X_t,Y_t)\,\mathrm{d}M_t - \Lambda\,\boldsymbol{\varphi}(X_t,Y_t)\,\mathrm{d}t, \quad \text{where } \Lambda = \begin{bmatrix} -\Lambda_{0,1} & \Lambda_{0,1}\\ \Lambda_{1,0} & -\Lambda_{1,0}\end{bmatrix}. \tag{1.3.12}$$

We remark that the debt process Y_t is constant and equal to $Y_{T\wedge\theta_k}$ over the interval $[T\wedge\theta_k, \widetilde{t}_{m,k})$.

By the Lagrange theorem, and since $\partial\varphi_j/\partial x \geq 1$ by (7.4.1), we obtain the value of the jump $\varphi_j(X_t, Y_{T\wedge\theta_k}) - \varphi_j(X_{t^-}, Y_{T\wedge\theta_k}) \leq X_t - X_{t^-} = -(Z_t - Z_{t^-})$. Then we take the expectation, noticing that again by (7.4.1), we have $\rho\,\boldsymbol{\varphi} - \mathrm{L}\boldsymbol{\varphi} - \Lambda\,\boldsymbol{\varphi} - \mathcal{J}\boldsymbol{\varphi} \geq 0$, where we have also considered the compensator found in (1.3.12), to eventually get

$$\begin{aligned}
\mathbb{E}\big[e^{-\rho\,\widetilde{t}_{m,k}}\,\boldsymbol{\varphi}(X_{\widetilde{t}^-_{m,k}}, Y_{T\wedge\theta_k})\big] \leq& \\
&\mathbb{E}\big[e^{-\rho\,(T\wedge\theta_k)}\,\boldsymbol{\varphi}(X_{T\wedge\theta_k}, Y_{T\wedge\theta_k})\big] - \mathbb{E}\bigg[\int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\,\mathrm{d}Z^c_t\bigg] \\
&- \mathbb{E}\bigg[\sum_{T\wedge\theta_k<t<\widetilde{t}_{m,k}} e^{-\rho\,t}\,(Z_t - Z_{t^-})\bigg] - \mathbb{E}\bigg[\int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\;\mathcal{J}\boldsymbol{\varphi}(X_t, Y_{T\wedge\theta_k})\,\mathrm{d}t\bigg],
\end{aligned}$$

where, by the localization of the process, the Brownian part disappears since it is a true martingale, and the same happens with the martingale in (1.3.12). Putting together Z^c with the jump part

$$\begin{aligned}
\mathbb{E}\big[e^{-\rho\,(T\wedge\theta_k)}\,\boldsymbol{\varphi}(X_{T\wedge\theta_k}, Y_{T\wedge\theta_k})\big] \geq& \\
&\mathbb{E}\bigg[\int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\,\mathrm{d}Z_t + \int_{T\wedge\theta_k}^{\widetilde{t}_{m,k}} e^{-\rho\,t}\;\mathcal{J}\boldsymbol{\varphi}(X_t, Y_{T\wedge\theta_k})\,\mathrm{d}t + e^{-\rho\,\widetilde{t}_{m,k}}\,\boldsymbol{\varphi}(X_{\widetilde{t}^-_{m,k}}, Y_{T\wedge\theta_k})\bigg].
\end{aligned}$$

Since $\mathcal{J}\boldsymbol{\varphi}(X_t, Y_{T\wedge\theta_k}) \geq 0$ and $\int_{T\wedge\theta_k}^{\widetilde{t}^-_{m,k}} e^{-\rho\,t}\,\mathrm{d}Z_t + e^{-\rho\,\widetilde{t}_{m,k}}\,\varphi_j(X_{\widetilde{t}^-_{m,k}}, Y_{T\wedge\theta_k}) \geq -P\,Y_{T\wedge\theta_k}$, i.e. the argument of the RHS is bounded from below, we take the limit for $m \to +\infty$ and by Fatou's lemma we obtain

$$\begin{aligned}
\mathbb{E}\big[e^{-\rho\,(T\wedge\theta_k)}\,\boldsymbol{\varphi}(X_{T\wedge\theta_k}, Y_{T\wedge\theta_k})\big] \geq& \mathbb{E}\bigg[\int_{T\wedge\theta_k}^{T\wedge\theta_{k+1}} e^{-\rho\,t}\,\mathrm{d}Z_t + \int_{T\wedge\theta_k}^{(T\wedge\theta_{k+1})^-} e^{-\rho\,t}\,\mathcal{J}\boldsymbol{\varphi}(X_t, Y_{T\wedge\theta_k})\,\mathrm{d}t \\
&+ e^{-\rho\,(T\wedge\theta_{k+1})}\,\boldsymbol{\varphi}(X_{(T\wedge\theta_{k+1})^-}, Y_{T\wedge\theta_k})\bigg].
\end{aligned} \tag{1.3.13}$$

Then we have

$$\mathbb{E}\big[e^{-\rho\,(T\wedge\theta_k)}\,\boldsymbol{\varphi}(X_{T\wedge\theta_k}, Y_{T\wedge\theta_k})\big] \geq \mathbb{E}\bigg[\int_{T\wedge\theta_k}^{T\wedge\theta_{k+1}} e^{-\rho\,t}\,dZ_t + e^{-\rho\,(T\wedge\theta_{k+1})}\,\boldsymbol{\varphi}(X_{T\wedge\theta_{k+1}}, Y_{T\wedge\theta_{k+1}})\bigg],$$

since, considering the upper boundary in the integration interval in (1.3.13), $\theta_{k+1}\wedge T$, it suffices to check the cases in which the debt offer arrival happens first, and the case in which the default event happens first. Consider first the case $\{\theta_{k+1} < T\}$, and we obtain that the expectation of the last two terms in (1.3.13) satisfy the inequality

$$\mathbb{E}\Big[\int_{\theta_k}^{\theta_{k+1}^-} e^{-\rho\, t}\,\mathcal{J}\boldsymbol{\varphi}(X_t, Y_{\theta_k})\,\mathrm{d}t + e^{-\rho\,\theta_{k+1}}\,\boldsymbol{\varphi}(X_{\theta_{k+1}^-}, Y_{\theta_k})\Big] = \mathbb{E}\big[e^{-\rho\,\theta_{k+1}}\,\boldsymbol{\varphi}(X_{\theta_{k+1}}, Y_{\theta_{k+1}})\big], \tag{1.3.14}$$

by the fact that $\int_0^t e^{-\rho\, s}\,\mathcal{J}\boldsymbol{\varphi}(X_s, Y_s)\,\mathrm{d}s$ is the compensator of the càdlàg pure jump process
$\sum_{k=0}^{N_t} e^{-\rho\,\theta_k}\big(\boldsymbol{\varphi}(X_{\theta_k}, Y_{\theta_k}) - \boldsymbol{\varphi}(X_{\theta_k^-}, Y_{\theta_k})\big)$, and therefore their difference is a martingale. Turning on the case $\{T \geq \theta_{k+1}\}$, we have $\varphi_j(X_T, Y_T) = -PY_T$, since $X_T \leq Y_T$, and hence we obtain a similar inequality as (1.3.14).

Iterating this procedure for all the indexes k such that $\theta_k < T$ we obtain

$$\varphi_j(x,y) \geq \mathbb{E}\Big[\int_0^{T^-} e^{-\rho\, t}\,\mathrm{d}Z_t + e^{-\rho\, T}\,\varphi_j(X_{T^-}, Y_T)\Big].$$

We remark that, by the finiteness of the intensity of the Poisson process, in finite intervals we have a finite number of jumps. □

1.3.4 Linear growth condition and continuity

We can now show that condition (1.3.4) is not only necessary but also sufficient to have that v is finite.

Corollary 1.3.4 *Assume (1.3.4), for all $(x,y) \in D$, $j \in \{0,1\}$, we have*

$$v_j(x,y) \leq E_j\, x - F_j\, y + G_j\,. \tag{1.3.15}$$

In particular the value function is finite.

Proof. Assume (1.3.9) is satisfied and let us consider the following functions

$$\begin{aligned}
\varphi_0(x,y) &= x - F_0\, y + G_0\,,\\
\varphi_1(x,y) &= \begin{cases} x - F_1\, y + G_1 & \text{if } \rho \geq b_1\,,\\ \frac{\Lambda_{1,0}}{\rho - b_1 + \Lambda_{1,0}} x - F_1\, y + G_1 & \text{if } \rho < b_1\,,\end{cases}
\end{aligned}$$

where (G_0, G_1) are large enough and the couple (F_0, F_1) is the unique solution to the system $(\rho + \Lambda_{j,1-j})F_j - \Lambda_{j,1-j}F_{1-j} - \overline{r}E_j - \lambda\int_{\mathbb{R}}[(E_j - F_j)\zeta]^+\,\nu(\mathrm{d}\zeta) = 0$, and $\overline{r}$ is the upperbound of the interest rate.

A tedious computation shows that this function is a supersolution. Therefore the conclusion comes directly by Proposition 1.3.3. □

We are now left with the study of the continuity of the value function.

Proposition 1.3.5 *The value functions $v_j(\cdot, y)$ are continuous on D and satisfy*

$$v_j(y^+, y) := \lim_{x \to y^+} v_j(x, y) = -P\, y\,,$$

Proof. a) For what concerns the continuity on the border, consider the auxiliary process $S^{j,x,y}$, see (1.2.1), and define the first hitting time of the lower barrier y, that is $\Theta_{j,x,y} := \inf\{t \geq 0 : S_t^{j,x,y} = y\}$. Notice that $S^{j,x,y}$ is dominated by $S^{j,x,0}$, i.e. a geometric Brownian motion, and define its hitting time on y as $\Theta_{j,x,0} := \inf\{t \geq 0 : S_t^{j,x,0} = y\}$. Moreover, define $\overline{\Theta}_{j,x,0} := \inf\{t \geq 0 : \overline{S}_t^{j,x,0} = y\}$ where $\overline{S}$ is the auxiliary process with constant drift b_1. Therefore we have the following inequalities

$$\Theta_{j,x,y} \leq \Theta_{j,x,0} \leq \overline{\Theta}_{j,x,0}, \quad \forall j \in \{0,1\}, x \geq y\,. \tag{1.3.16}$$

Fix some $\gamma > 0$ such that $y < x < y+\gamma$ and denote $\overline{\Theta}^{\gamma}_{j,x,y} = \inf\{t \geq 0 \,:\, \overline{S}_t^{j,x,y} = y+\gamma\}$, hence, by Proposition 3.2 in [104, Chapter VII], we have

$$\mathbb{P}\left(\overline{\Theta}_{j,x,0} > \overline{\Theta}^{\gamma}_{j,x,0}\right) = \frac{s(x) - s(y)}{s(y+\gamma) - s(y)}\,,$$

with θ_1 the time of the first arrival of a investment/divestment offer where s is a scale function of the process $\overline{S}^{j,x,0}$.

$$\mathbb{P}\left(\overline{\Theta}_{j,x,0} > \overline{\Theta}^{\gamma}_{j,x,0}\right) \to 0\,, \quad \text{for } x \to y\,.$$

For $y+\gamma$ we have the previous inequality on the hitting times $\overline{\Theta}^{\gamma}_{j,x,0} \leq \Theta^{\gamma}_{j,x,0} \leq \Theta^{\gamma}_{j,x,y}$. Therefore, combing with (1.3.16) we obtain $\mathbb{P}(\Theta_{j,x,y} > \Theta^{\gamma}_{j,x,y}) \leq \mathbb{P}(\overline{\Theta}_{j,x,0} > \overline{\Theta}^{\gamma}_{j,x,0})$, hence

$$\mathbb{P}\left(\Theta_{j,x,y} > \Theta^{\gamma}_{j,x,y}\right) \to 0, \quad \text{for } x \to y\,. \tag{1.3.17}$$

Moreover, we have that the first opportunity arrival time is less than the hitting barrier for $\overline{S}$ with probability given by

$$\begin{aligned}\mathbb{P}\left((\theta_1 \wedge \tau_1) < \overline{\Theta}_{j,x,0} < \overline{\Theta}^{\gamma}_{j,x,0}\right) &= \mathbb{E}\left[\mathbb{1}_{(\theta_1 \wedge \tau_1) < \overline{\Theta}_{j,x,0} < \overline{\Theta}^{\gamma}_{j,x,0}}\right] \\ &= \mathbb{E}\left[\mathbb{E}[\mathbb{1}_{(\theta_1 \wedge \tau_1) < \overline{\Theta}_{j,x,0}} \mid \mathcal{F}_{\overline{\Theta}_{j,x,0}}]\, \mathbb{1}_{\overline{\Theta}_{j,x,0} < \overline{\Theta}^{\gamma}_{j,x,0}}\right]\end{aligned}$$

$$= \mathbb{E}\left[\left(1-e^{-\lambda\,\overline{\Theta}_{j,x,0}}\right)\,\left(1-e^{-\Lambda\,\overline{\Theta}_{j,x,0}}\right)\,\mathbb{1}_{\overline{\Theta}_{j,x,0}<\overline{\Theta}^{\gamma}_{j,x,0}}\right],$$

and by equation 2.1.4.(1) in [24, page 622], since the integrand is bounded, we can take the limit $x \to y$ inside the integral, and the integral tends to zero, therefore

$$\mathbb{P}\big((\theta_1 \wedge \tau_1) < \overline{\Theta}_{j,x,0} < \overline{\Theta}^{\gamma}_{j,x,0}\big) \to 0\,, \quad \text{for } x \to y\,. \tag{1.3.18}$$

Let $\alpha = (Z, \{\pi_k\}_{k\geq 1})$ be an arbitrary policy, and denote $\eta = T^{j,x,y,\alpha} \wedge \overline{\Theta}^{\gamma}_{j,x,y} \wedge \Theta_{j,x,y} \wedge \theta_1 \wedge \tau_1$. Notice that $\eta \leq \Theta^{\gamma}_{j,x,y}$. Taking the limit $x \to y$, for $t \leq \eta$ no offers are arrived a.s., hence $X^{j,x,y,\alpha}_t \leq S^{j,x,y,\alpha}_t$ and also $T^{j,x,y} \leq \Theta_{j,x,y}$, which together with (1.3.18) implies $\lim_{x\to y} \eta = T^{j,x,y,\alpha} \wedge \overline{\Theta}^{\gamma}_{j,x,y}$ a.s..

Then, considering x approaching y, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T^-} e^{-\rho\,t}\,\mathrm{d}Z_t\right] \\
&\quad = \lim_{x\to y}\left(\mathbb{E}\left[\int_0^{\eta^-} e^{-\rho\,t}\,\mathrm{d}Z_t\right] + \mathbb{E}\left[\mathbb{1}_{T>\eta}\int_{\eta}^{T^-} e^{-\rho\,t}\,\mathrm{d}Z_t\right]\right) \\
&\quad \leq \lim_{x\to y}\left(\mathbb{E}\left[Z_{\eta^-}\right] + \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{T>\eta}\int_{\eta}^{T^-} e^{-\rho\,t}\,\mathrm{d}Z_t\,\middle|\,\mathcal{F}_{\overline{\Theta}^{\gamma}_{j,x,y}}\right]\right]\right) \\
&\quad \leq \lim_{x\to y}\left(\mathbb{E}\left[\overline{S}^{j,x,0}_{\eta} - y\right] + \mathbb{E}\left[\mathbb{1}_{T>\overline{\Theta}^{\gamma}_{j,x,y}}\,\mathbb{E}\left[\int_{\overline{\Theta}^{\gamma}_{j,x,y}}^{T^-} e^{-\rho\,t}\,\mathrm{d}Z_t\,\middle|\,\mathcal{F}_{\overline{\Theta}^{\gamma}_{j,x,y}}\right]\right]\right) \\
&\quad \leq \lim_{x\to y}\left(\mathbb{E}[\overline{S}^{j,x,0}_{\eta} - y] + \mathbb{E}\left[\mathbb{1}_{T>\overline{\Theta}^{\gamma}_{j,x,y}}\,e^{-\rho\,\overline{\Theta}^{\gamma}_{j,x,y}}\left(v_j\left(X^{i,x,y}_{\overline{\Theta}^{\gamma}_{j,x,y}}, Y^{i,x,y}_{\overline{\Theta}^{\gamma}_{j,x,y}}\right) + P\,Y^{i,x,y}_{\overline{\Theta}^{\gamma}_{j,x,y}}\right)\right]\right).
\end{aligned} \tag{1.3.19}$$

Now, since v_j is non-decreasing w.r.t. the first component and $Y_t = y$ is fixed for $t \leq \overline{\Theta}^{\gamma}_{j,x,y}$ and $x \to y$, since $\overline{\Theta}^{\gamma}_{j,x,y} < \theta_1$ a.s., we have $v_j\left(X^{i,x,y}_{\overline{\Theta}^{\gamma}_{j,x,y}}, Y^{i,x,y}_{\overline{\Theta}^{\gamma}_{j,x,y}}\right) \leq v_j(y+\gamma, y)$. Then by (1.3.17) and (1.3.19), we obtain

$$\begin{aligned}
0 &\leq \mathbb{E}\left[\int_0^{T^-} e^{-\rho\,t}\,\mathrm{d}Z_t\right] \\
&\leq \lim_{x\to y}\left(\mathbb{E}\left[\sup_{0\leq t\leq\overline{\Theta}^{\gamma}_{j,x,y}} \overline{S}^{j,x,0}_t - y\right] + \big(v_j(y+\gamma, y) + P\,y\big)\,\mathbb{P}\big(\overline{\Theta}_{j,x,y} > \overline{\Theta}^{\gamma}_{j,x,y}\big)\right) = 0\,.
\end{aligned}$$

So, by the properties of hitting times of geometric Brownian motion,

$$\mathbb{E}\left[-P\,Y_T\,e^{-\rho T}\right] \leq -P\,\mathbb{E}\left[Y_{\overline{\Theta}_{j,x,0}}e^{-\rho\,\overline{\Theta}_{j,x,0}}\right] \to -P\,y,$$

and we obtain

$$
\begin{aligned}
-P\,y \le v_j(x,y) &\le \sup_{\alpha\in\mathcal{A}} \mathbb{E}\left[\int_0^{T^-} e^{-\rho\, t}\, \mathrm{d}Z_t - P\, Y_T\, e^{-\rho\, T}\right] \\
&\le \lim_{x\to y}\Bigg(\mathbb{E}\left[\sup_{0\le t\le \overline{\Theta}^{\gamma}_{j,x,y}} \overline{S}^{j,x,0}_t - y\right] + \big(v_j(y+\gamma,y) + P\,y\big)\,\mathbb{P}\Big(\overline{\Theta}_{j,x,y} > \overline{\Theta}_{j,x,y}\Big) \\
&\quad - P\,\mathbb{E}\left[Y_{\overline{\Theta}_{j,x,0}} e^{-\rho\,\overline{\Theta}_{j,x,0}}\right]\Bigg) \quad = \quad -P\,y\,,
\end{aligned}
$$

and finally we proved $v_j(y^+,y) = -P\,y$.

b) Now we prove the continuity of the value function $v_j(\cdot,y)$ w.r.t to the first component on $(y,+\infty)$. Consider the auxiliary process S solution to (1.2.1), and define $T^{\widehat{\alpha},\gamma}$ as the hitting time of the point $x+\gamma$ for S, i.e.

$$
T^{\alpha,\gamma} := \inf\{t \ge 0 :\, S^{(j,x,y,\alpha)}_t \ge x+\gamma\}\,.
$$

This time depends on the chosen strategy. Therefore consider $\widehat{\alpha}$ the admissible strategy in which dividends are not paid, and new debt facilities offers are not accepted, i.e. $X^{(j,x,y,\widehat{\alpha})}_t = S^{(j,x,y,\widehat{\alpha})}_t$ and $Y^{(j,x,y,\widehat{\alpha})}_t = y$ up to τ_1. Since this strategy is not said to be the optimal, by the dynamic programming principle, we have

$$
v_j(x,y) \ge \mathbb{E}\Big[e^{-\rho\,(T^{\widehat{\alpha}}\wedge T^{\widehat{\alpha},\gamma}\wedge\tau_1)}\, v_j\big(X_{T^{\widehat{\alpha}}\wedge T^{\widehat{\alpha},\gamma}\wedge\tau_1}, Y_{T^{\widehat{\alpha}}\wedge T^{\widehat{\alpha},\gamma}\wedge\tau_1}\big)\Big]\,,
$$

where we recall that $T^{\widehat{\alpha}}$ is the bankruptcy time and comes from (1.1.3). Notice that the term inside the expectation can be rewritten as

$$
e^{-\rho\, T^{\widehat{\alpha},\gamma}\wedge\tau_1}\, v_j(X_{T^{\widehat{\alpha},\gamma}\wedge\tau_1}, Y_{T^{\widehat{\alpha},\gamma}\wedge\tau_1})\, \mathbb{1}_{T^{\widehat{\alpha},\gamma}\wedge\tau_1 < T^{\widehat{\alpha}}} - P\, Y_{T^{\widehat{\alpha}}}\, e^{-\rho\, T^{\widehat{\alpha}}}\, \mathbb{1}_{T^{\widehat{\alpha},\gamma}\wedge\tau_1 \ge T^{\widehat{\alpha}}}\,,
$$

moreover $X_{T^{\widehat{\alpha},\gamma}} \ge x+\gamma$. Therefore, since v_j is non-decreasing with respect to its first component we obtain

$$
\begin{aligned}
&v_j(x{+}\gamma,y) - v_j(x,y) \\
&\quad\le \mathbb{E}\big[\big(1-e^{-\rho\, T^{\widehat{\alpha},\gamma}}\big) v_j(x+\gamma,y)\, \mathbb{1}_{T^{\widehat{\alpha},\gamma} < T^{\widehat{\alpha}}\wedge\tau_1} + v_j(x+\gamma,y)\, \mathbb{1}_{T^{\widehat{\alpha},\gamma}\wedge\tau_1 \ge T^{\widehat{\alpha}}}\big] \\
&\qquad + \mathbb{E}\big[P\, Y_{T^{\widehat{\alpha}}}\, e^{-\rho\, T^{\widehat{\alpha}}}\, \mathbb{1}_{T^{\widehat{\alpha},\gamma}\wedge\tau_1 \ge T^{\widehat{\alpha}}}\big] + \mathbb{E}\big[\big(v_j(x+\gamma,y) - e^{-\rho\,\tau_1} v_j(X_{\tau_1},y)\big)\, \mathbb{1}_{\tau_1 \le T^{\widehat{\alpha},\gamma}\wedge T^{\widehat{\alpha}}}\big] \\
&\quad\le v_j(x+\gamma,y)\, \Big(1 - \mathbb{E}\Big[e^{-\rho\, T^{\widehat{\alpha},\gamma}}\Big]\Big) + \big(v_j(x+\gamma,y) + Py\big)\, \mathbb{P}(T^{\widehat{\alpha},\gamma}\wedge\tau_1 \ge T^{\widehat{\alpha}}) \\
&\qquad + \big(v_j(x+\gamma,y) + Py\big)\, \mathbb{P}(\tau_1 \le T^{\widehat{\alpha},\gamma}\wedge T^{\widehat{\alpha}})\,,
\end{aligned}
$$

where the last term is obtained using $v_j(x,y) \geq Py$. Now consider a fixed upper barrier in $x+\gamma_0$ and consider another one in $x+\gamma$ with $\gamma \leq \gamma_0$, then we have

$$\begin{aligned}
&v_j(x+\gamma,y) - v_j(x,y) \\
&\quad \leq v_j(x+\gamma_0,y)\left(1-\mathbb{E}\left[e^{-\rho\,(T^{\widehat{\alpha},\gamma}\wedge\tau_1)}\right]\right) + \left(v_j(x+\gamma_0,y)+Py\right)\mathbb{P}(T^{\widehat{\alpha}}\wedge\tau_1 \leq T^{\widehat{\alpha},\gamma}).
\end{aligned}$$

By Proposition 3.2 in [104, Chapter VII, Section 3], and using the same arguments as for the right continuity on the border, we have

$$\begin{aligned}
\mathbb{P}\left(T^{\widehat{\alpha},\gamma} \geq T^{\widehat{\alpha}}\right) &\;\to\; 0, &&\text{for } \gamma \to 0,\\
\mathbb{P}\left(T^{\widehat{\alpha},\gamma} \geq \tau_1\right) &\;\to\; 0, &&\text{for } \gamma \to 0, \text{ and}\\
\mathbb{E}\left[e^{-\rho\,(T^{\widehat{\alpha},\gamma}\wedge\tau_1)}\right] &\;\to\; 1, &&\text{for } \gamma \to 0.
\end{aligned}$$

By the linear growth condition on v_j in Corollary 1.3.4, we reach the right-continuity also in the interior of the domain, letting $\gamma \to 0$. A similar argument gives the left-continuity. □

1.3.5 Unique viscosity solution

Theorem 1.3.6 *The value functions $v_j(\cdot, y)$ are continuous on D and constitute the unique viscosity solution to the system of variational inequalities:*

$$\min\left\{\rho\, v_j(x,y) - \mathcal{L}_j v_j(x,y) - \mathcal{J} v_j(x,y) - \mathcal{G}_j v_j(x,y)\,;\; \frac{\partial v_j(x,y)}{\partial x} - 1\right\} = 0, \quad x > y, \tag{1.3.20}$$

with linear growth condition in both x and y given by equation (1.3.15), and the boundary condition

$$v_j(y^+,y) := \lim_{x\to y^+} v_j(x,y) = -P\,y\,.$$

The proof is based on the following lemmas.

Lemma 1.3.3 *The value function v defined by (1.1.5) is a viscosity supersolution to the system of variational inequalities (1.2.3).*

Proof. Let $(\overline{x},\overline{y}) \in D$ and consider a C^2-test function (φ_0,φ_1) such that $\varphi_j \leq v_j$ and $v_{\overline{j}}(\overline{x},\overline{y}) = \varphi_{\overline{j}}(\overline{x},\overline{y})$. Assume w.l.o.g. that $(\overline{x},\overline{y})$ is a minimum for $v_j - \varphi_j$ on the neighborhood $B^+_\epsilon(\overline{x},\overline{y}) := \{(x,y) \in \mathbb{R}^+ \times \mathbb{R}^+ : |x-\overline{x}| \leq \epsilon, |y-\overline{y}| \leq \epsilon, x > y\}$ for

$j \in \{0,1\}$. We have to prove that

$$\min\Big\{\rho\,\varphi_{\bar{j}}(\overline{x},\overline{y}) - \mathcal{L}_j\varphi_{\bar{j}}(\overline{x},\overline{y}) - \mathcal{J}\varphi_{\bar{j}}(\overline{x},\overline{y}) - \mathcal{G}_j\varphi_{\bar{j}}(\overline{x},\overline{y})\,;\ \frac{\partial\varphi_{\bar{j}}(\overline{x},\overline{y})}{\partial x} - 1\Big\} \geq 0\,. \quad (1.3.21)$$

Notice that the manager at each time is allowed to pay everything in dividends, therefore we have the lower bound $\varphi_{\bar{j}}(\overline{x},\overline{y}) = v_{\bar{j}}(\overline{x},\overline{y}) \geq \overline{x} - \overline{y} - P\,\overline{y}$. Consider the admissible strategy in which at time 0 a quantity $0 < \delta < (\overline{x}-\overline{y}) \wedge \epsilon$ is paid in dividends, therefore, by the fact that the value function dominates all strategies, we have

$$\varphi_{\bar{j}}(\overline{x},\overline{y}) = v_{\bar{j}}(\overline{x},\overline{y}) \geq v_{\bar{j}}(\overline{x}-\delta,\overline{y}) + \delta\,,$$

by the supersolution condition, we have $\varphi_{\bar{j}}(\overline{x},\overline{y}) - \varphi_{\bar{j}}(\overline{x}-\delta,\overline{y}) \geq \delta$. Then dividing by δ, consider the limit $\delta \to 0$, to get that that the left derivative of $\varphi_{\bar{j}}$ with respect to the first component has to be greater or equal than 1, but since $\varphi_j \in C^2$, the left derivative is equal to right derivative, hence we obtain

$$\frac{\partial\varphi_{\bar{j}}(\overline{x},\overline{y})}{\partial x} - 1 \geq 0\,,$$

so we have just to prove the inequality for the first member in (1.3.21), i.e.

$$\rho\,\varphi_{\bar{j}}(\overline{x},\overline{y}) - \mathcal{L}_j\varphi_{\bar{j}}(\overline{x},\overline{y}) - \mathcal{J}\varphi_{\bar{j}}(\overline{x},\overline{y}) - \mathcal{G}_j\varphi_{\bar{j}}(\overline{x},\overline{y}) \geq 0\,.$$

Now define the exit time

$$t_\epsilon = \inf\{t \geq 0 : (M_t^{\bar{j}}, X_t^{\overline{x}}, Y_t^{\overline{y}}) \notin \{\bar{j}\} \times B_\epsilon^+(\overline{x},\overline{y})\}\,,$$

and consider a general stopping time $h > 0$, a strategy consisting in no dividend payment up to time h, and again by the dynamic programming principle we have

$$\begin{aligned}
\varphi_{\bar{j}}(\overline{x},\overline{y}) = v_{\bar{j}}(\overline{x},\overline{y}) \geq \mathbb{E}\big[&v_{\bar{j}}(S_{h\wedge t_\epsilon},\overline{y})\,\mathbb{1}_{h\wedge t_\epsilon<\theta\wedge\tau\wedge T} \\
&+ \max\{v_{\bar{j}}(S_\theta + \overline{y}\zeta, \overline{y}(\zeta+1)), v_{\bar{j}}(S_\theta,\overline{y})\}\,\mathbb{1}_{\theta<h\wedge t_\epsilon\wedge\tau\wedge T} \\
&+ v_{1-\bar{j}}(S_\tau,\overline{y})\,\mathbb{1}_{\tau<h\wedge t_\epsilon\wedge\theta\wedge T} - P\,\overline{y}\,e^{-\rho T}\,\mathbb{1}_{T<h\wedge t_\epsilon\wedge\tau\wedge\theta}\big] \\
\geq \mathbb{E}\big[&\varphi_{\bar{j}}(S_{h\wedge t_\epsilon},\overline{y})\,\mathbb{1}_{h\wedge t_\epsilon<\theta\wedge\tau\wedge T} \\
&+ \max\{\varphi_{\bar{j}}(S_\theta + \overline{y}\zeta, \overline{y}(\zeta+1)), \varphi_{\bar{j}}(S_\theta,\overline{y})\}\,\mathbb{1}_{\theta<h\wedge t_\epsilon\wedge\tau\wedge T} \\
&+ \varphi_{1-\bar{j}}(S_\tau,\overline{y})\,\mathbb{1}_{\tau<h\wedge t_\epsilon\wedge\theta\wedge T} - P\,\overline{y}\,e^{-\rho T}\,\mathbb{1}_{T<h\wedge t_\epsilon\wedge\tau\wedge\theta}\big],
\end{aligned} \quad (1.3.22)$$

where θ and τ are the times respectively of first opportunity arrival and the first switch in the regime and S is the auxiliary process given by (1.2.1). Apply Itô formula to $e^{-\rho t}\varphi_{\bar{j}}(S_t, \bar{y})$ for times t such that

$$0 \ \leq \ t \ < \ \gamma_\epsilon := t_\epsilon \wedge h \wedge \tau \wedge \theta \, ,$$

and take the expectation to get

$$\mathbb{E}\left[e^{-\rho \gamma_\epsilon}\, \varphi_{\bar{j}}(S_{\gamma_\epsilon}, \bar{y})\right] = \varphi_{\bar{j}}(\bar{x}, \bar{y}) + \mathbb{E}\left[\int_0^{\gamma_\epsilon} e^{-\rho\, t}\left(-\rho\, \varphi_{\bar{j}} + \mathcal{L}_j \varphi_{\bar{j}}\right)(S_t, \bar{y})\,\mathrm{d}t\right],$$

where we remark that $X_{\gamma_\epsilon^-} = S_{\gamma_\epsilon}$. Combining with inequality (1.3.22), and noticing that $\mathbb{E}[e^{-\rho\, t}\, \varphi_j(X_t, Y_t)] \leq \mathbb{E}[\varphi_j(X_t, Y_t)]$, we obtain

$$\begin{aligned} 0 \geq \mathbb{E}\Bigg[\int_0^{\gamma_\epsilon} e^{-\rho\, t}\big(-\rho\, \varphi_{\bar{j}} + \mathcal{L}_j \varphi_{\bar{j}}\big)(S_t, \bar{y})\,\mathrm{d}t\Bigg] + \mathbb{E}\Big[e^{-\rho\, \theta}\Big(\max\{\varphi_{\bar{j}}(S_\theta + \bar{y}\zeta, \bar{y}(\zeta+1)), \varphi_{\bar{j}}(S_\theta, \bar{y})\} \\ - \varphi_{\bar{j}}(S_{\theta^-}, \bar{y})\Big)\, \mathbb{1}_{\theta < \gamma_\epsilon \wedge \tau}\Big] + \mathbb{E}\Big[e^{-\rho\, \tau}\Big(\varphi_{1-\bar{j}}(S_\tau, \bar{y}) - \varphi_{\bar{j}}(S_{\tau^-}, \bar{y})\Big)\, \mathbb{1}_{\tau < \gamma_\epsilon \wedge \theta}\Big] \\ \geq \mathbb{E}\Bigg[\int_0^{\gamma_\epsilon} e^{-\rho\, t}\left(-\rho\, \varphi_{\bar{j}} + \mathcal{L}_j \varphi_{\bar{j}} + \mathcal{J}\varphi_{\bar{j}} + \mathcal{G}_j \varphi_{\bar{j}}\right)(S_t, \bar{y})\,\mathrm{d}t\Bigg]. \end{aligned} \tag{1.3.23}$$

By definition of the exit time t_ϵ, we see that the integrand part of (1.3.23) is bounded, take the limit $\epsilon \to 0$, and obtain the supersolution property by the mean value theorem. □

Lemma 1.3.4 *The value function v_j defined by (1.1.5), for $j \in \{0, 1\}$, is a viscosity subsolution to the system of variational inequalities (1.2.3).*

Proof. To prove the subsolution property we proceed by contradiction, i.e. we assume that there exists a constant $\epsilon > 0$, a point $(\bar{j}, \bar{x}, \bar{y}) \in \{0, 1\} \times \mathbb{R}^+ \times \mathbb{R}^+$, with $(\bar{x}, \bar{y}) \in D$, a C^2-function φ_j, $j \in \{0.1\}$, such that $(\varphi_{\bar{j}} - v_{\bar{j}})(\bar{x}, \bar{y}) = 0$ and $\varphi_j \geq v_j$ on the neighborhood $B_{(\bar{x},\bar{y})}(\epsilon) := \{(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+ \ : \ |x - \bar{x}| \leq \epsilon, |y - \bar{y}| \leq \epsilon, (x, y) \in D\}$ for $j \in \{0, 1\}$, and $\eta > 0$ such that for all $(x, y) \in B_{(\bar{x},\bar{y})}(\epsilon)$ we have

$$\left(\rho\, \varphi_{\bar{j}} - \mathcal{L}_j \varphi_{\bar{j}} - \mathcal{J}\varphi_{\bar{j}} - \mathcal{G}_j \varphi_{\bar{j}}\right)(x, y) > \eta \, , \tag{1.3.24}$$

$$\frac{\partial \varphi_{\bar{j}}(x, y)}{\partial x} - 1 > \eta \, . \tag{1.3.25}$$

Then consider the exit time from the ball

$$t_\epsilon := \inf\left\{t \geq 0 \ : \ (M_t^{\bar{j}}, X_t^{\bar{x}}, Y_t^{\bar{y}}) \notin \{\bar{j}\} \times B_{(\bar{x},\bar{y})}(\epsilon)\right\},$$

define the time $\gamma_\epsilon := t_\epsilon \wedge \theta \wedge T = t_\epsilon \wedge \theta$, apply Itô formula to $e^{-\rho t}\,\boldsymbol{\varphi}(X_t, \overline{y})$ for $t \in [0, t_\epsilon)$, and take the expectation to get

$$\begin{aligned}\mathbb{E}\left[e^{-\rho\,\gamma_\epsilon}\boldsymbol{\varphi}(X_{\gamma_\epsilon}, \overline{y})\right] =& \boldsymbol{\varphi}(\overline{x}, \overline{y}) + \mathbb{E}\bigg[\int_0^{\gamma_\epsilon} e^{-\rho\,t}\big(-\rho\,\boldsymbol{\varphi} + \mathrm{L}\boldsymbol{\varphi} + \Lambda\boldsymbol{\varphi}\big)(S_t, \overline{y})\,\mathrm{d}t \\ & - \int_0^{\gamma_\epsilon} e^{-\rho\,t}\,\frac{\partial\varphi}{\partial x}(X_t, \overline{y})\,\mathrm{d}Z_t^c - \sum_{0<t<\gamma_\epsilon} e^{-\rho\,t}\left(\boldsymbol{\varphi}(X_t, \overline{y}) - \boldsymbol{\varphi}(X_{t^-}, \overline{y})\right)\bigg],\end{aligned}$$

since

$$\int \mathrm{J}_2\,\boldsymbol{\varphi}(X_t, y_k)\,\mathrm{d}M_t - \Lambda\,\boldsymbol{\varphi}(X_t, y_k)\,\mathrm{d}t, \quad \text{where } \Lambda = \begin{bmatrix} -\Lambda_{0,1} & \Lambda_{0,1} \\ \Lambda_{1,0} & -\Lambda_{1,0} \end{bmatrix},$$

is a martingale. From inequality (1.3.24) and (1.3.25) we obtain

$$\begin{aligned}\boldsymbol{\varphi}(\overline{x}, \overline{y}) &\geq \mathbb{E}\bigg[e^{-\rho\,\gamma_\epsilon}\,\boldsymbol{\varphi}(X_{\gamma_\epsilon}, \overline{y}) + \int_0^{\gamma_\epsilon} e^{-\rho\,t}\left(\eta + \mathcal{J}\boldsymbol{\varphi}(X_t, \overline{y})\right)\mathrm{d}t + \int_0^{\gamma_\epsilon}(1+\eta)\,e^{-\rho\,t}\,\mathrm{d}Z_t\bigg] \\ &\geq \mathbb{E}\left[e^{-\rho\,\gamma_\epsilon}\boldsymbol{\varphi}(X_{\gamma_\epsilon}, \overline{y}) + \int_0^{\gamma_\epsilon} e^{-\rho\,t}\,\mathrm{d}Z_t\right] + \eta\,\mathbb{E}\left[\int_0^{\gamma_\epsilon} e^{-\rho\,t}\,\mathrm{d}t + \int_0^{\gamma_\epsilon} e^{-\rho\,t}\,\mathrm{d}Z_t\right] \\ &\geq \mathbb{E}\left[e^{-\rho\,\gamma_\epsilon}\boldsymbol{v}(X_{\gamma_\epsilon}, \overline{y}) + \int_0^{\gamma_\epsilon} e^{-\rho\,t}\,\mathrm{d}Z_t\right] + \eta\,\mathbb{E}\left[\frac{1 - e^{-\rho\,(t_\epsilon \wedge \theta)}}{\rho} + \int_0^{\gamma_\epsilon} e^{-\rho\,t}\,\mathrm{d}Z_t\right],\end{aligned} \tag{1.3.26}$$

coming from the fact that $\mathcal{J}\boldsymbol{\varphi}(X_s, Y_s)$ is non-negative, and take C as the multiplication term of η in (1.3.26). We easily deduce that C is strictly positive, since the two terms can not be both null. By the dynamic programming principle, we obtain

$$\boldsymbol{\varphi}(\overline{x}, \overline{y}) \geq \boldsymbol{v}(\overline{x}, \overline{y}) + \eta\,C,$$

and hence the contradiction. □

The following Lemma is the key to see the uniqueness of the solution, since it gives a strict super-solution dominating the value function.

Lemma 1.3.5 *Let $\big(w_j(\cdot, y)\big)_{j\in\{0,1\}}$ be a continuous viscosity super-solution to the system of variational inequalities (1.3.20) on D, and define the following function*

$$h_j := A + E_j(x + \mathrm{e})\,\log(x + \mathrm{e}) + y\,\log(y + \mathrm{e})\,, \tag{1.3.27}$$

with $(E_0, E_1) = \left(1, \max\left\{1, \frac{\Lambda_{1,0}}{\rho + \Lambda_{1,0} - b_1}\right\}\right)$ and A_j positive constant large enough.

Then, for all $\gamma \in (0, 1)$, $w_j^\gamma := (1-\gamma)\,w_j + \gamma\,h_j$ is a strict supersolution to the

HJB equation (1.2.3), i.e. there exists some $\delta > 0$ such that

$$\min\left\{\rho\, w_j^\gamma - \mathcal{L}_j w_j^\gamma - \mathcal{J} w_j^\gamma - \mathcal{G}_j w_j^\gamma\,;\ \frac{\partial w_j^\gamma(x,y)}{\partial x} - 1\right\} \geq \delta\,. \tag{1.3.28}$$

Proof. Let u_j and w_j continuous viscosity subsolution and supersolution, respectively, to the system of variational inequalities, for $j \in \{0,1\}$ and with starting asset and debt value $(x,y) \in D$. Assume that they satisfy the boundary condition $u_j(y^+, y) \leq w_j(y^+, y)$, with the linear growth condition

$$|u_j(x,y)| + |w_j(x,y)| \leq C_1 + C_2\, x, \quad \forall j \in \{0,1\},$$

for some positive constants C_1 and C_2. We want to prove that in the interior the same inequality as the boundary condition holds, i.e. $u_j(x,y) \leq w_j(x,y)$, for $j \in \{0,1\}$, $(x,y) \in D$.
Step 1: Consider the strict supersolution h_j to (1.2.3) as in equation (1.3.27), and its linear combination with the supersolution w_j, $w_j^\gamma := (1-\gamma)\, w_j + \gamma\, h_j$, for $j \in \{0,1\}$. Notice that h_j dominates w_j.

Starting from the second term in (1.3.28), we have

$$\frac{\partial w_j^\gamma}{\partial x} = (1-\gamma)\,\frac{\partial w_j}{\partial x} + \gamma\,\frac{\partial h_j}{\partial x} \geq 1 + \delta\,, \tag{1.3.29}$$

since the first term of the RHS is greater or equal to 1 since w_j is supersolution, while

$$\frac{\partial h_j}{\partial x} = E_j\left[1 + \log(x + \mathrm{e})\right]$$

is strictly greater than 1, and we have the required inequality for the second term in (1.3.28). For the first term in (1.3.28), we have

$$\rho\, h_j - \mathcal{L}_j h_j - \mathcal{J} h_j - \mathcal{G}_j h_j = \left[E_j(\rho + \Lambda_{j,1-j} - b_j) - \Lambda_{j,1-j} E_{1-j}\right]\, x\, \log(x + \mathrm{e}) + o(x)\,,$$

where the dominant term is always non-negative thanks to (1.3.9). By straightforward calculation, for $x \in (y, \infty)$, we have

$$\rho\, h_j - \mathcal{L}_j h_j - \mathcal{J} h_j - \mathcal{G}_j h_j \geq \delta > 0\,, \tag{1.3.30}$$

by taking A large enough. Therefore h_j is a supersolution dominating $v(x,y)$ for $|(x,y)| \to \infty$.

Combining (1.3.29) with (1.3.30), for $j \in \{0,1\}$, we have inequality (1.3.28).

Step 2: In order to prove the comparison principle, it suffices to show that for all $\gamma \in (0,1)$:

$$\max_{j\in\{0,1\}} \sup_D \big(u_j(x,y) - w_j^\gamma(x,y)\big) \le 0\,,$$

since the required result is obtained by letting γ to 0. We argue by contradiction and suppose that there exists some $\gamma \in (0,1)$ and $\bar{j} \in \{0,1\}$ such that

$$\vartheta := \max_{j\in\{0,1\}} \sup_D \big(u_j(x,y) - w_j^\gamma(x,y)\big) = \sup_D \big(u_{\bar{j}}(x,y) - w_{\bar{j}}^\gamma(x,y)\big) > 0\,.$$

Notice that $u_j(x,y) - w_j^\gamma(x,y)$ tends to minus infinity when x tends to infinity. We also have $\lim_{x\to y^+}\big(u_j(x,y) - w_j^\gamma(x,y)\big) \le \gamma\left(\lim_{x\to y^+} w_j(x,y) - h_j(y,y)\right)$. Hence, by the continuity of the functions u_j and w_j^γ, there exists $x_0 \in (y,\infty)$ such that

$$\vartheta = u_j(x_0,y) - w_j^\gamma(x_0,y)\,.$$

For any $\epsilon > 0$, we consider the functions

$$\Phi_\epsilon(x,x') = u_j(x,y) - w_j^\gamma(x',y) - \phi_\epsilon(x,x')\,,$$
$$\phi_\epsilon(x,x') = \frac{1}{4}|x-x_0|^4 + \frac{1}{2\epsilon}|x-x'|^2\,,$$

for all $x, x' \in (y,\infty)$. By standard arguments in comparison principle, the function Φ_ϵ attains a maximum in $(x_\epsilon, x'_\epsilon) \in (y,\infty)^2$, which converges (up to a subsequence) to (x_0,x_0) when ϵ goes to zero. Moreover,

$$\lim_{\epsilon\to 0} \frac{|x_\epsilon - x'_\epsilon|^2}{\epsilon} = 0\,.$$

Applying Theorem 3.2 in [41], we get the existence of $M_\epsilon, M'_\epsilon \in \mathbb{R}$ such that:

$$(p_\epsilon, M_\epsilon) \in J^{2,+}u_j(x_\epsilon)\,,$$
$$(p'_\epsilon, M'_\epsilon) \in J^{2,+}w_j^\gamma(x'_\epsilon)\,,$$

and,

$$\begin{bmatrix} M_\epsilon & 0 \\ 0 & -M'_\epsilon \end{bmatrix} \le D^2\phi_\epsilon(x_\epsilon,x'_\epsilon) + \epsilon\left(D^2\phi_\epsilon(x_\epsilon,x'_\epsilon)\right)^2\,, \tag{1.3.31}$$

where I_2 is the 2×2-identity matrix, and

$$\begin{aligned} p_\epsilon &= D_x\phi_\epsilon(x_\epsilon, x'_\epsilon) = \frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) + (x_\epsilon - x_0)^3, \\ p'_\epsilon &= D_{x'}\phi_\epsilon(x_\epsilon, x'_\epsilon) = \frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon)\,, \\ D^2\phi_\epsilon(x_\epsilon, x'_\epsilon) &= \begin{bmatrix} 3(x_\epsilon - x_0)^2 + \frac{1}{\epsilon} & -\frac{1}{\epsilon} \\ -\frac{1}{\epsilon} & \frac{1}{\epsilon} \end{bmatrix}. \end{aligned}$$

By writing the viscosity subsolution property of u_j and the strict supersolution property (1.3.28) of w_j^γ, we have the following inequalities:

$$\begin{aligned} \min\Big\{ &\rho\, u_j(x_\epsilon, y) - (b_j\, x_\epsilon - r\, y)\left(\frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) + (x_\epsilon - x_0)^3\right) - \frac{1}{2}\sigma^2\, x_\epsilon^2\, M_\epsilon \\ &- \mathcal{J} u_j(x_\epsilon, y) - \mathcal{G}_j u_j(x_\epsilon, y)\,;\ \frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) + (x_\epsilon - x_0)^3 - 1\Big\} \le 0\,, \qquad (1.3.32) \\ \min\Big\{ &\rho\, w_j^\gamma(x'_\epsilon, y) - (b_j\, x'_\epsilon - r\, y)\,\frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) - \frac{1}{2}\sigma^2\, x'^2_\epsilon\, M'_\epsilon \\ &- \mathcal{J} w_j^\gamma(x'_\epsilon, y) - \mathcal{G}_j w_j^\gamma(x'_\epsilon, y)\,;\ \frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) - 1\Big\} \ge \delta\,. \qquad (1.3.33) \end{aligned}$$

We then distinguish the following two cases depending on (1.3.32):

- Case 1: $\frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) + (x_\epsilon - x_0)^3 \le 1$. Notice that by (1.3.33), we have

  $$\frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) \ge 1 + \delta\,,$$

  which implies

  $$(x_\epsilon - x_0)^3 \le -\delta\,.$$

  By sending ϵ to zero, we obtain a contradiction.

- Case 2:

  $$\begin{aligned} \rho\, u_j(x_\epsilon, y) - (b_j\, x_\epsilon - r\, y)\left(\frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) + (x_\epsilon - x_0)^3\right) \\ - \frac{1}{2}\sigma^2\, x_\epsilon^2\, M_\epsilon - \mathcal{J} u_j(x_\epsilon, y) - \mathcal{G}_j u_j(x_\epsilon, y) \le 0\,. \qquad (1.3.34) \end{aligned}$$

  From (1.3.33), we have

  $$\rho\, w_j^\gamma(x'_\epsilon, y) - (b_j\, x'_\epsilon - r\, y)\,\frac{1}{\epsilon}\,(x_\epsilon - x'_\epsilon) - \frac{1}{2}\sigma^2\, x'^2_\epsilon\, M'_\epsilon - \mathcal{J} w_j^\gamma(x'_\epsilon, y) - \mathcal{G}_j w_j^\gamma(x'_\epsilon, y) \ge \delta, \qquad (1.3.35)$$

  Using (1.3.31), we obtain an upper bound for $\frac{1}{\epsilon}\,(x_\epsilon^2\, M_\epsilon - x'^2_\epsilon\, M'_\epsilon)$. Combing it with (1.3.34)−(1.3.35), yields to an upper bound for $\rho\left(u_j(x_\epsilon, y) - w_j^\gamma(x'_\epsilon, y)\right)$, which goes to $-\delta$ when we send ϵ to zero.

Using the continuity of u_j and w_j^γ, we obtain the required contradiction $\rho\,\vartheta \leq \delta < 0$.

□

1.4 Picard iteration scheme

The aim of this section is to provide approximation schemes to the HJB equation (1.2.3) through a Picard's type sequence, starting from an exact solution of a simplified problem. Let us start with the definition of the Picard sequence from the probabilistic point of view, and then move to the HJB sequence which will be proven to be satisfied by the Picard sequence, see Lemma 1.4.1.

Definition 1.4.1 (Picard iteration) Let $N \in \mathbb{N}$ and $j \in \{0,1\}$, we define the N-th element of the sequence of Picard iteration $\{v_j^{(N)}\}_{N\in\mathbb{N}}$ as the value function $v_j^{(N)}$ corresponding to the maximization problem as in Section 1.1, except that the underlying asset-debt dynamics is getting N arrival-event, where by *arrival-event* we mean a switch in the business cycle's state or an investment/divestment opportunity arrival, whether it is accepted or not, i.e.

$$v_j^{(N)}(x,y) = \sup_{\alpha_N \in \mathcal{A}_N} \mathbb{E}\left[\int_0^{T^-} e^{-\rho\,t}\,\mathrm{d}Z_t - P\,Y_T\,e^{-\rho\,T}\right], \quad x \geq y \geq \underline{y}\,, \tag{1.4.1}$$

where $\mathcal{A}_N$ is the set of admissible strategies for a manager able to accept only the opportunities among the first N arrival-events, and the dynamics of X^{α_N} and Y^{α_N} come from

$$\begin{cases} \mathrm{d}X_t = (b_{\widetilde{M}_t^{(N)}}\,X_t - r(Y_T)\,Y_t)\,\mathrm{d}t + \sigma\,X_t\,\mathrm{d}W_t - \mathrm{d}Z_t + \mathrm{d}Y_t \quad \forall t \geq 0 \\ X_0 = x \end{cases}$$

with $\widetilde{M}_t^{(N)}$, $Y_t^{\alpha_N}$ coincide with M, Y up to the N-arrival event and constant afterwards.

The particular case $N = 0$, i.e. $v_j^{(0)}$, is the first element of the Picard iteration and its value function corresponds to the maximization problem for an underlying asset-debt structure with no arrival events; in other words, the value function of an optimal dividend problem for a diffusion model.

Definition 1.4.2 (HJB sequence) We define the sequence of solutions to the HJB sequence $\{v_j^{(N)}\}_{N\in\mathbb{N}}$ starting at $v_j^{(0)}$ solution of the linear PDE of the second order:

$$\begin{aligned}&\min\big\{\rho\, v_j^{(0)}(x,y) - \mathcal{L}_j v_j^{(0)}(x,y)\,;\ \partial_x v_j^{(0)}(x,y) - 1\big\} = 0, \quad \text{for } (x,y)\in D\,, \\ &v_j^{(0)}(y,y) = -P\,y\,,\end{aligned} \tag{1.4.2}$$

by the following recurrence relation for the next elements of the sequence $N\in\mathbb{N}$

$$\begin{aligned}\min\Bigg\{&\rho\, v_j^{(N+1)}(x,y) - \mathcal{L}_j v_j^{(N+1)}(x,y) - \lambda\int_{\mathbb{R}}\big[\max\big(v_j^{(N)}(x+y\,\zeta, y\,(1+\zeta)), v_j^{(N)}(x,y)\big) \\ &- v_j^{(N+1)}(x,y)\big]\,\nu(\mathrm{d}\zeta) - \Lambda_{j,1-j}\big(v_{1-j}^{(N)}(x,y) - v_j^{(N+1)}(x,y)\big)\,;\ \partial_x v_j^{(N+1)}(x,y) - 1\Bigg\} = 0\,,\end{aligned} \tag{1.4.3}$$

with initial condition $v_j^{(N+1)}(y,y) = -P\,y$.

First of all, we would like to prove that $\{v^{(N)}\}_{N\in\mathbb{N}}$ converges in Cauchy sense, and then it converges to the value function v defined by HJB (1.2.3), and therefore v can be obtained through the method of successive approximation starting from the exact solution of (1.4.2).

In the next subsection we will prove that indeed the Picard iteration defined in Definition 1.4.1 is solution to the HJB sequence 1.4.2, see Lemma 1.4.1.

1.4.1 Convergence of approximation sequence

Lemma 1.4.1 *Let $N\in\mathbb{N}$ and $(x,y)\in D$, and assume (1.3.4) is satisfied*

- *Classic supersolution: if $\{\varphi_j^{(N+1)}\}_{j\in\{0,1\}}\in C^{2,0}(D)$ is such that*

$$\begin{aligned}\varphi_j^{(N+1)}(y^+,y) &:= \lim_{x\downarrow y}\varphi_j^{(N+1)}(x,y) \geq -P\,y\,, \\ \varphi_j^{(N+1)}(x,y) &= -\,P\,y, \quad \textit{for } x<y\,,\end{aligned}$$

and satisfies

$$\min\Big\{(\rho-\mathcal{L}_j+\lambda)\varphi_j^{(N+1)}(x,y)-\lambda\left(\int_{\mathbb{R}}\max\{v_j^{(N)}(x+y\zeta,y(1+\zeta)),v_j^{(N)}(x,y)\}\nu(\mathrm{d}\zeta)\right)$$
$$-\Lambda_{j,1-j}\big(v_{1-j}^{(N)}(x,y)-\varphi_j^{(N+1)}(x,y)\big)\,;\ \partial_x\varphi_j^{(N+1)}(x,y)-1\Big\}\geq 0\,,$$

then $\varphi_j \geq v_j^{(N+1)}$.

- *Linear growth: we have*

$$v_j^{(N+1)}(x,y)\leq E_j\,x-F_j\,y+G_j\,.$$

- *Continuity: the value functions* $v_j^{(N+1)}(\cdot,y)$ *are continuous on* D *and satisfy*

$$v_j^{(N+1)}(y^+,y):=\lim_{x\to y^+}v_j^{(N+1)}(x,y)=-P\,y\,.$$

- *Dividend area: there exists* $k^*\in\mathbb{R}^+$ *such that* $x_y^*<y+k^*$, *such that*

$$v_j^{(N+1)}(x,y)=x-x_y^*+v_j^{(N+1)}(x_y^*,y),\quad \textit{for each } x>x_y^* \textit{ and each } N\in\mathbb{N}.$$

- *Viscosity characterization: the value functions* $v_j^{(N+1)}$, *defined by* (1.4.1), *constitute the unique viscosity solution to the system of variational inequalities* (1.4.3).

Proof. The proof is analogous to the results in Proposition 1.3.3, and 1.3.5, Corollary 1.3.4, and Theorem 1.3.6, and it is therefore omitted. □

Let us now show that the solutions to the sequence of HJB obtained as a linearization to equation (1.2.3) is a contracting sequence.

Lemma 1.4.2 (Compactness) *Let* $D^*:=\{(x,y)\in D\ :\ \ x\leq y+k^*\}\subset D$, *for* $N\in\mathbb{N}$ *and* k^* *defined as in Lemma 1.4.1, there exists a positive real constant* $C^{*,N}$ *such that*

$$\big\|v_j^{(N+1)}(x,y)-v_j^{(N)}(x,y)\big\|_{L^\infty(D)}=\big\|v_j^{(N+1)}(x,y)-v_j^{(N)}(x,y)\big\|_{L^\infty(D^*)}=C^{*,N}.$$

Proof. For $x\geq y+k^*$ and each $N\in\mathbb{N}$, the value functions $v_j^{(N+1)}$ and $v_j^{(N)}$ are just vertical translations of each other, since $\forall\, N \in \mathbb{N}$ and $\forall\, x \geq y + k^*$

$$v_j^{(N)}(x,y) = v_j^{(N)}(y+k^*,y) + x - y - k^*,$$

therefore there exist a positive constant $C^{*,N}$ such that

$$\left\| v_j^{(N+1)}(x,y) - v_j^{(N)}(x,y) \right\|_{L^\infty(D \setminus D^*)} = |v_j^{(N+1)}(y+k^*,y) - v_j^{(N)}(y+k^*,y)| =: C^{*,N}.$$

□

Lemma 1.4.2 proves that

$$\operatorname*{arg\,max}_{(x,y)\in D} |v_j^{(N+1)}(x,y) - v_j^{(N)}(x,y)| \in D^*;$$

this compactness result will be used in the next proposition, stating that the sequence of approximations is a contraction.

Proposition 1.4.3 *Let* $\mathbf{v} := [v_0, v_1]$*, then the sequence* $\{\mathbf{v}^{(N)}\}_{N\in\mathbb{N}}$ *given by* (1.4.3) *is a contracting sequence in the* $L^\infty(\mathbb{R}^2)$*-norm, that is there exists a constant* $C \in (0,1)$ *such that*

$$\|\mathbf{v}^{(N+1)}(x,y) - \mathbf{v}^{(N)}(x,y)\|_{L^\infty(D)} \leq C \; \|\mathbf{v}^{(N)}(x,y) - \mathbf{v}^{(N-1)}(x,y)\|_{L^\infty(D)}\,, \quad \textit{for } N \in \mathbb{N}, \tag{1.4.4}$$

defining $v_j^{(-1)}(x,y) := x - y - P\,y$ *for* $j \in \{0,1\}$.

Proof. Notice that by the compactness lemma 1.4.2 and the continuity property of $v_j^{(N)}(x,y)$ by Lemma 1.4.1 we can consider $(x,y) \in D^*$ instead than in the whole unbounded $\overline{D}$. Then let φ be stopping time, by the DPP for (1.4.1), we have

$$\begin{aligned} v_j^{(N+1)}(x,y) = \sup_{\alpha\in\mathcal{A}_{N+1}(0,\varphi)} \mathbb{E}\bigg[& \int_0^{T\wedge\varphi\wedge\theta_1\wedge\tau_1} e^{-\rho\, t}\, \mathrm{d}Z_t - P\,y\, e^{-\rho\, T^\alpha}\, \mathbb{1}_{T<\varphi\wedge\theta_1\wedge\tau_1} \\ & + \Big(v_j^{(N)}(X^\alpha_{\theta_1^-} + y\,\zeta, y\,(1+\zeta)) - v_j^{(N)}(X^\alpha_{\theta_1^-}, y) \Big)_+ \mathbb{1}_{\theta_1<\varphi\wedge T\wedge\tau_1} \\ & + v_j^{(N+1)}(X^\alpha_\varphi, y)\, \mathbb{1}_{\varphi<T\wedge\theta_1\wedge\tau_1} + v_{1-j}^{(N)}(X^\alpha_{\tau_1}, y)\, \mathbb{1}_{\tau_1<\theta_1\wedge\varphi\wedge T} \bigg]. \end{aligned}$$

Consider the optimal strategy $\alpha_{N+1,*}$ (denoted as α_* to simplify the notations where no ambiguity occurs) in the interval $[0,\varphi]$ for a manager able to accept the opportunities among the first $N+1$ arrival events, and apply it to a problem allowing the acceptance of only the opportunities among the first N arrival events, with the difference that the $(N+1)$-st opportunity will be rejected. Notice that this strategy $\alpha_{N+1,*}$ belongs to the admissible strategies $\mathcal{A}_N$, and that if the default time for $v_j^{(N+1)}$ is less or equal to θ_{N+1}, then in any case it coincides to the default time in the case in which the $(N+1)$-st opportunity has to be rejected. We have

$$\begin{aligned}v_j^{(N)}(x,y) \geq \mathbb{E}\Bigg[&\int_0^{T\wedge\varphi\wedge\theta_1\wedge\tau_1} e^{-\rho\, t}\,\mathrm{d}Z_t - P\, y\, e^{-\rho T}\,\mathbb{1}_{T<\varphi\wedge\theta_1\wedge\tau_1} + v_j^{(N)}(X_\varphi^{\alpha_*},y)\,\mathbb{1}_{\varphi<T\wedge\theta_1\wedge\tau_1}\\
&+\Big(v_j^{(N-1)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta, y\,(1+\zeta)) - v_j^{(N-1)}(X_{\theta_1^-}^{\alpha_*},y)\Big)\,\mathbb{1}_{\theta_1<\varphi\wedge T\wedge\tau_1}\\
&\mathbb{1}_{\left\{v_j^{(N)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta))\geq v_j^{(N)}(X_{\theta_1^-}^{\alpha_*},y)\right\}} + v_{1-j}^{(N-1)}(X_{\tau_1}^{\alpha},y)\,\mathbb{1}_{\tau_1<\theta_1\wedge\varphi\wedge T}\Bigg].\end{aligned}$$

Therefore taking into consideration the difference between $v_j^{(N+1)}$ and $v_j^{(N)}$ we have

$$\begin{aligned}&v_j^{(N+1)}(x,y) - v_j^{(N)}(x,y) \leq\\
&\mathbb{E}\Bigg[\Big(v_j^{(N+1)}(X_\varphi^{\alpha_*},y) - v_j^{(N)}(X_\varphi^{\alpha_*},y)\Big)\,\mathbb{1}_{\varphi<T\wedge\theta_1\wedge\tau_1} + \mathbb{1}_{\left\{v_j^{(N)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta))\geq v_j^{(N)}(X_{\theta_1^-}^{\alpha_*},y)\right\}}\,\mathbb{1}_{\theta_1<\varphi\wedge T\wedge\tau_1}\\
&\times\Big(v_j^{(N)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta)) - v_j^{(N)}(X_{\theta_1^-}^{\alpha_*},y) - v_j^{(N-1)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta)) + v_j^{(N-1)}(X_{\theta_1^-}^{\alpha_*},y)\Big)\\
&+\Big(v_{1-j}^{(N)}(X_{\tau_1}^{\alpha_*},y) - v_{1-j}^{(N-1)}(X_{\tau_1}^{\alpha_*},y)\Big)\,\mathbb{1}_{\tau_1<\theta_1\wedge\varphi\wedge T}\Bigg]. \qquad (1.4.5)\end{aligned}$$

Let T^* be the default time for an enterprise with starting asset value x_y^* and paying no dividends, see Lemma 1.4.1. Then, for $\epsilon>0$ and $\varphi=T^*+\epsilon$, we have that $T<T^*$ for $x\in[y,x_y^*]$, and therefore equation (1.4.5) becomes

$$\begin{aligned}v_j^{(N+1)}(x,y) - v_j^{(N)}(x,y) \leq \mathbb{E}\Bigg[&\Bigg\{\Big(v_j^{(N)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta)) - v_j^{(N-1)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta))\Big)\\
&-\Big(v_j^{(N)}(X_{\theta_1^-}^{\alpha_*},y) - v_j^{(N-1)}(X_{\theta_1^-}^{\alpha_*},y)\Big)\Bigg\}\,\mathbb{1}_{\left\{v_j^{(N)}(X_{\theta_1^-}^{\alpha_*}+y\,\zeta,y\,(1+\zeta))\geq v_j^{(N)}(X_{\theta_1^-}^{\alpha_*},y)\right\}}\,\mathbb{1}_{\theta_1<T\wedge\tau_1}\\
&+\big(v_{1-j}^{(N)}(X_{\tau_1}^{\alpha_*},y) - v_{1-j}^{(N-1)}(X_{\tau_1}^{\alpha_*},y)\big)\,\mathbb{1}_{\tau_1<\theta_1\wedge T}\Bigg]. \qquad (1.4.6)\end{aligned}$$

A similar argument works also considering the difference $v_j^{(N)} - v_j^{(N+1)}$: apply to both the optimization problem for $v_j^{(N)}$ and the optimization problem for $v_j^{(N+1)}$ the optimal strategy for the manager able to accept only the opportunities among the first N arrival events. Therefore we can take the absolute values on both sides of (1.4.6), and recalling that $\mathbf{v} := [v_0, v_1]$, by Hölder inequality and triangle inequality, we obtain

$$
\begin{aligned}
&|\mathbf{v}^{(N+1)}(x,y) - \mathbf{v}^{(N)}(x,y)| \\
&\leq \|\mathbb{1}_{\theta_1 \wedge \tau_1 < T}\|_{L^1(\mathrm{d}\mathbb{P})} \cdot \Big\| \mathbb{P}(\theta_1 < \tau_1) \left| \mathbf{v}^{(N)}\big(X_{\theta_1}^{\alpha_{N+1,*}}, Y_{\theta_1}^{\alpha_{N+1,*}}\big) - \mathbf{v}^{(N-1)}\big(X_{\theta_1}^{\alpha_{N+1,*}}, Y_{\theta_1}^{\alpha_{N+1,*}}\big) \right| \\
&\qquad + \mathrm{J}_2\, \mathbb{P}(\tau_1 < \theta_1) \left| \mathbf{v}^{(N)}\big(X_{\tau_1}^{\alpha_{N+1,*}}, Y_{\tau_1}^{\alpha_{N+1,*}}\big) - \mathbf{v}^{(N-1)}\big(X_{\tau_1}^{\alpha_{N+1,*}}, Y_{\tau_1}^{\alpha_{N+1,*}}\big) \right| \Big\|_{L^\infty(\mathrm{d}\mathbb{P})},
\end{aligned}
$$

for every $(x,y) \in D^*$, where by $|\cdot|$ we mean the distance in an L^2-space. Equivalently we may write

$$
\begin{aligned}
&\left|\left(\mathbf{v}^{(N+1)}(x,y) - \mathbf{v}^{(N)}(x,y)\right)\right| \\
&\leq \mathbb{E}_{(x,y)} \left[\sup_{\substack{\theta_1 > 0 \\ \omega \in \mathcal{F}_{\theta_1}}} \left\{ \left| \mathbf{v}^{(N)}\left(X_{\theta_1}^{\alpha_{N+1,*}}, Y_{\theta_1}^{\alpha_{N+1,*}}\right) - \mathbf{v}^{(N-1)}\left(X_{\theta_1}^{\alpha_{N+1,*}}, Y_{\theta_1}^{\alpha_{N+1,*}}\right) \right| \right\} \right] \mathbb{P}\left(\theta_1 \wedge \tau_1 < T\right).
\end{aligned}
$$

For the expectation term we may take the supreme of its argument

$$
\begin{aligned}
\sup_{\theta_1 \geq 0,\, \zeta \in (-1,\infty),\, \omega \in \mathcal{F}_{\theta_1}^-} &\Big\{ \Big| \mathbf{v}^{(N)}\left(X_{\theta_1^-}^{Z^{\alpha^*}} + y\, \pi_1^{\alpha^*} \zeta, y(1 + \pi_1^{\alpha^*} \zeta) \right) \\
&\qquad - \mathbf{v}^{(N-1)}\left(X_{\theta_1^-}^{Z^{\alpha^*}} + y\, \pi_1^{\alpha^*} \zeta, y\,(1 + \pi_1^{\alpha^*} \zeta) \right) \Big| \Big\},
\end{aligned}
$$

and then $\forall (x,y) \in D^*$

$$
\left|\left(\mathbf{v}^{(N+1)}(x,y) - \mathbf{v}^{(N)}(x,y)\right)\right| \leq \|\mathbf{v}^{(N)}(x,y) - \mathbf{v}^{(N-1)}(x,y)\|_{L^\infty(D^*)}\, \mathbb{P}\left(\theta_1 \wedge \tau_1 < T\right). \tag{1.4.7}
$$

Notice that the second term of the RHS in (1.4.7) is equal to $1 - \mathbb{P}(\theta_1 > T) \cdot \mathbb{P}(\tau_1 > T)$; therefore it is enough to show that $\mathbb{P}(\theta_1 > T)$ is non null and a similar argument works also for the other term, since they are both exponential random variables. We have that θ_1 is just an $\mathcal{F}_t$-stopping time independent on asset and debt values, while T depends on the optimal strategy and both on x and y. We consider a process $\widehat{X}$ with the following dynamics

$$
\begin{cases}
\mathrm{d}\widehat{X}_t = b_j\, \widehat{X}_t\, \mathrm{d}t + \sigma\, \widehat{X}_t\, \mathrm{d}W_t\,, \quad t > 0\,, \\
\widehat{X}_0 = k^* + y\,,
\end{cases}
$$

see Lemma 1.4.1 for the definition of k^*. It is easy to see that $\widehat{X}_t$ dominates X_t. We have that, for $t \in \mathbb{R}^+$, $\widehat{X}_t = k^* \exp\{(b_j - \sigma^2/2)\, t + \sigma\, W_t\}$. Then, define $\widehat{T} := \inf\{t \geq$

$0 \ : \ 0 \leq \widehat{X}_t\} = \inf\{t \geq 0 \ : \ \widehat{X}_t = 0\}$, i.e. the first hitting time for a geometric Brownian motion, independent on both x and y, and therefore, by [24, 9.1.2.2 Part II], we have

$$\mathbb{P}(\theta_1 > \widehat{T}) = 1 - \mathbb{P}\left(\inf_{0\leq s\leq\theta_1} \widehat{X}_s > 0\right) = 0\,.$$

Hence taking the supremum of (1.4.7) over the couples $(x, y) \in D^*$, we obtain inequality (1.4.4). □

In the following proposition we prove that the sequence $\{v_j^{(N)}\}_{N\in\mathbb{N}}$ converges to the solution to HJB equation (1.2.3).

Proposition 1.4.4 *The sequence $\{v_j^{(N)}\}_{N\in\mathbb{N}}$ solution to (1.4.2) and (1.4.3) converges to the values function defined by equation (1.1.5).*

Proof. Proposition 1.4.3 guarantees the convergence of $\{v_j^{(N)}\}_{N\in\mathbb{N}}$ towards a solution in viscosity sense of HJB (1.4.3), the limit of $v_j^{(N)}$ clearly satisfies HJB (1.2.3), and then, by uniqueness result of Theorem 1.3.6, $v_j^{(N)}$ defined by (1.4.1) converges to the value function defined by (1.1.5). □

1.4.2 Exact solution for the approximated sequence

In this subsection we give the exact formulae for the starting function to the sequence of Picard iterations given by equation (1.4.2), and its following elements given by Definition 1.4.1.

Proposition 1.4.5 *The solutions to (1.4.2) and (1.4.3), for $N \in \mathbb{N}_+$, read as*

$$v_j^{(0)}(x,y) = \begin{cases} x^{-\widehat{A}_j^0}\left[C_1^0\,\Phi\Big(\widehat{A}_j^0,\ \widehat{B}_j^0;\ -\dfrac{2\,r\,y}{\sigma^2\,x}\Big) + C_2^0\,\Psi\Big(\widehat{A}_j^0,\ \widehat{B}_j^0;\ -\dfrac{2\,r\,y}{\sigma^2\,x}\Big)\right], & x < x_j^{0,*}(y), \\[2ex] x - x_j^{0,*}(y) + \big(x_j^{0,*}(y)\big)^{-\widehat{A}_j^0}\left[C_1^0\,\Phi\Big(\widehat{A}_j^0,\ \widehat{B}_j^0;\ -\dfrac{2\,r\,y}{\sigma^2\,x_j^{0,*}(y)}\Big)\right. \\ \qquad\qquad \left. +C_2^0\,\Psi\Big(\widehat{A}_j^0,\ \widehat{B}_j^0;\ -\dfrac{2\,r\,y}{\sigma^2\,x_j^{0,*}(y)}\Big)\right], & x \geq x_j^{0,*}(y), \end{cases}$$

$$v_j^{(N+1)}(x,y) = \begin{cases} C_1^{N+1}\,\Phi_j(x,y) + C_2^{N+1}\,\Psi_j(x,y) + \Theta_j^{(N)}(x,y), & x < x_j^{N+1,*}(y), \\[2ex] x - x_j^{N+1,*}(y) + C_1^{N+1}\,\Phi_j\big(x_j^{N+1,*}(y), y\big) \\ \qquad +C_2^{N+1}\,\Psi_j\big(x_j^{N+1,*}(y), y\big) + \Theta_j^{(N)}\big(x_j^{N+1,*}(y), y\big), & x \geq x_j^{N+1,*}(y), \end{cases}$$

where we have defined the parameter $\overline{\rho} := \rho + \lambda + \Lambda_{j,1-j}$,

$$\widehat{A}_j^0 := \frac{b_j}{\sigma^2} - \frac{1}{2} + \frac{1}{2}\sqrt{\Big(1 - \frac{2\,b_j}{\sigma^2}\Big)^2 + \frac{8\,\rho}{\sigma^2}}, \quad \widehat{A}_j := \frac{b_j}{\sigma^2} - \frac{1}{2} + \frac{1}{2}\sqrt{\Big(1 - \frac{2\,b_j}{\sigma^2}\Big)^2 + \frac{8\,\overline{\rho}}{\sigma^2}},$$
$$\widehat{B}_j^0 := 1 + \sqrt{\Big(1 - \frac{2\,b_j}{\sigma^2}\Big)^2 + \frac{8\,\rho}{\sigma^2}}, \qquad \widehat{B}_j := 1 + \sqrt{\Big(1 - \frac{2\,b_j}{\sigma^2}\Big)^2 + \frac{8\,\overline{\rho}}{\sigma^2}},$$

and the following auxiliary functions

$$\begin{aligned}
\Phi_j(x,y) &:= x^{-\widehat{A}_j}\, \Phi\Big(\widehat{A}_j,\ \widehat{B}_j;\ -\frac{2\,r\,y}{\sigma^2\,x}\Big),\\
\Psi_j(x,y) &:= x^{-\widehat{A}_j}\, \Phi\Big(\widehat{A}_j,\ \widehat{B}_j;\ -\frac{2\,r\,y}{\sigma^2\,x}\Big),\\
\Upsilon_j^{(N)}(x,y) &:= -\lambda \int_{\mathbb{R}} \Big(v_j^{(N)}(x + y\,\zeta, y\,(1+\zeta)) - v_j^{(N)}(x,y)\Big)_+ \nu(\mathrm{d}\zeta)\\
&\qquad\qquad - \Lambda_{j,1-j} v_{1-j}^{(N)}(x,y) - \lambda v_j^{(N)}(x,y),\\
\Theta_j^{(N)}(x,y) &:= \int_y^\infty \Upsilon_j^{(N)}(z,y) \left[\Big(\Phi_j(x,y)\Big)^{-2} z^{-\frac{2b_j}{\sigma^2}} e^{-\frac{2ry}{\sigma^2 z}}\right] \mathrm{d}z\,.
\end{aligned}$$

Moreover, for $N \in \mathbb{N}$, $x_j^{N,*}(y) \in [y, +\infty)$, C_1^N *and* C_2^N *are given by imposing the initial condition and the smooth-fit condition*

$$\begin{cases} v_j^{(N)}(y,y) = -P \\ \partial_x v_j^{(N)}\big(x_j^{N,*}(y), y\big) = 1 \\ \partial_{xx}^2 v_j^{(N)}\big(x_j^{N,*}(y), y\big) = 0 \end{cases}$$

and the functions Φ *and* Ψ *are the confluent hypergeometric functions of the first and second type respectively.*

The proof is based on tedious but fairly standard calculations consisting in substituting all of the defined terms in equation (1.4.2).

Chapter 2

Comparative statics

2.1 Dimension reduction

This section deals with an important invariance in our model. We suppose that the interest rate r is constant and independent from the debt level y. This hypothesis could be easily justified remarking that the bankruptcy is declared at the first time when the asset process X reaches the debt level Y. Due to the continuity of the evolution of X and its cash equivalence, the manager could refund debt holders entirely. In this case the penalty P is justified by the breach of perennial loan agreement.

Our main result is to show that the value function v that depends both on the asset value x and the debt value y could be written using the debt y as unit of account. We will introduce an auxiliary value function $\widetilde{v}$ depending only on the asset value written on the debt $u := x/y$.

Proposition 2.1.1 (Debt as unit of account) *Let $U_t := X_t/Y_t$ and $\widetilde{Z}_t := \int_0^t Y_{s^-}^{-1}\,\mathrm{d}Z_s$ be respectively the asset value X and the cumulated dividend process Z written using the debt Y as unit of account, the related SDE reads*

$$\mathrm{d}U_t = (b_j\,U_t - r)\;\mathrm{d}t + \sigma\,U_t\,\mathrm{d}W_t - \mathrm{d}\widetilde{Z}_t + \mathrm{d}\widetilde{Y}_t\,, \tag{2.1.1}$$

where the effect of a change of the debt is given by $\mathrm{d}\widetilde{Y}_{\theta_i} = -\pi_i \frac{\zeta_i}{\zeta_i+1}\,(U_{\theta_i} - 1)$. Define

$$\widetilde{v}_j(u) := \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,u)}\left[\int_0^{T^-} e^{-\rho\,t}\prod_{k=1}^{\infty}(1+\zeta_k\,\pi_k\,\mathbb{1}_{\theta_k<t})\;\mathrm{d}\widetilde{Z}_t - Pe^{-\rho\,T}\prod_{k=1}^{\infty}(1+\zeta_k\,\pi_k\,\mathbb{1}_{\theta_k\le T})\right] \tag{2.1.2}$$

Then $\widetilde{v}_j(x/y) := \frac{1}{y}v_j(x,y)$ on $\{0,1\}\times[1,\infty)$. Moreover, we have the following dynamic *programming principle*

$$\widetilde{v}_j(u) := \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,u)}\Bigg[\int_0^{T^-\wedge\eta\wedge\theta_1} e^{-\rho t}\,\mathrm{d}\widetilde{Z}_t - Pe^{-\rho T}\,\mathbb{1}_{T<\eta\wedge\theta_1} + e^{-\rho\eta}\,\widetilde{v}_{M_\eta}(U_\eta)\,\mathbb{1}_{\eta<T\wedge\theta_1}$$
$$+ (1+\pi_1\zeta_1)\,\widetilde{v}_{M_{\theta_1}}\left(U_{\theta_1} - \pi_1\frac{\zeta_1}{\zeta_1+1}(U_{\theta_1}-1)\right)\mathbb{1}_{\theta_1<T\wedge\eta}\Bigg], \quad (2.1.3)$$

and $\widetilde{v}$ *is the unique viscosity solution of the system of variational inequalities*

$$\min\left\{\rho\widetilde{v}_j - \widetilde{\mathcal{L}}_j\widetilde{v}_j - \mathcal{G}_j\widetilde{v}_j - \widetilde{\mathcal{J}}\widetilde{v}_j(u)\,;\ \widetilde{v}_j' - 1\right\} = 0\,, \quad (2.1.4)$$

with $\widetilde{v}_j(1) = -P$ *where*

$$\begin{aligned}
\widetilde{\mathcal{L}}_j\widetilde{v}_j(u) &= \big(b_j\,u - r\big)\,\widetilde{v}_j'(u) + \frac{1}{2}\,\sigma^2\,u^2\,\widetilde{v}_j''(u)\,,\\
\mathcal{G}_j\widetilde{v}_j(u) &= \Lambda_{j,1-j}\,\big(\widetilde{v}_{1-j}(u) - \widetilde{v}_j(u)\big),\\
\widetilde{\mathcal{J}}\widetilde{v}_j(u) &= \lambda\int_{-1}^{\infty}\left((\zeta+1)\,\widetilde{v}_j\left(\frac{u-1}{\zeta+1}+1\right) - \widetilde{v}_j(u)\right)_+ \nu(\mathrm{d}\zeta)\,.
\end{aligned}$$

Proof. First, the SDE (2.1.1) is obtained by a direct computation of $U_t := X_t/Y_t$. We then remark that T coincides with $\inf\{t \mid U_t \le 1\}$. Since $\widetilde{Z}_t := \int_0^t Y_{s^-}^{-1} dZ_s$, we have $d\widetilde{Z}_t = dZ_t/Y_{t^-}$. Using the definition of value function (1.1.5) and dividing by y we found easily the relation between the two value function.

Using Theorem 1.3.6, we can deduce that $\{\widetilde{v}_j\}_{j=0,1}$ are the unique continuous viscosity solution of the system of inequalities (2.1.4).

We now turn to the dynamic programming principle, we write (2.1.2) in the following way

$$\begin{aligned}
\widetilde{v}_j(u) &= \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,u)}\Bigg[\int_0^{T^-\wedge\theta_1} e^{-\rho t}\,\mathrm{d}\widetilde{Z}_t - P\,e^{-\rho T} Y, \mathbb{1}_{T<\theta_1}\\
&\quad + \int_{T^-\wedge\theta_1}^{T^-} e^{-\rho t}\prod_{k=1}^{\infty}(1+\zeta_k\,\pi_k\,\mathbb{1}_{\theta_k<t})\,\mathrm{d}\widetilde{Z}_t - P\,e^{-\rho T}\,\mathbb{1}_{T\ge\theta_1}\prod_{k=1}^{\infty}(1+\zeta_k\,\pi_k\,\mathbb{1}_{\theta_k\le T})\Bigg]\\
&= \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,u)}\Bigg[\int_0^{T^-\wedge\theta_1} e^{-\rho t}\,\mathrm{d}\widetilde{Z}_t - Pe^{-\rho T}\mathbb{1}_{T<\theta_1} + (1+\pi_1\zeta_1\,\mathbb{1}_{T\ge\theta_1})\times\\
&\quad \times\left\{\int_{\theta_1}^{T^-} e^{-\rho t}\prod_{k=2}^{\infty}(1+\zeta_k\,\pi_k\,\mathbb{1}_{\theta_k<t})\,\mathrm{d}\widetilde{Z}_t - Pe^{-\rho T}\prod_{k=2}^{\infty}(1+\zeta_k\,\pi_k\,\mathbb{1}_{\theta_k\le T})\right\}\Bigg]\\
&= \sup_{\alpha\in\mathcal{A}} \mathbb{E}^{(j,u)}\Bigg[\int_0^{T^-\wedge\theta_1} e^{-\rho t}\,\mathrm{d}\widetilde{Z}_t - Pe^{-\rho T}\mathbb{1}_{T<\theta_1}
\end{aligned}$$

$$+(1+\pi_1\zeta_1)\,\mathbb{1}_{T\geq\theta_1}\widetilde{v}_{M_{\theta_1}}\left(U_{\theta_1^-}-\pi_1\frac{\zeta_1}{\zeta_1+1}\left(U_{\theta_1}-1\right)\right)\Bigg],$$

we consider the original one (7.3.1) stopped at $\eta\wedge\theta_1$ where η is a stopping time.

□

2.2 Sensitivity analysis

Using the reduced HJB equation (2.1.4), we perform a comparative statics. We first consider a single regime case and we fix the following parameters, in accord with Hugonnier et al. [68].

b	σ	λ	ρ	r
0.0824	0.2886	3	0.12	0.042

The law of proposal debt is $\nu(d\zeta)=\frac{1}{2}\delta_{(0.10)}(\zeta)+\frac{1}{2}\delta_{(-0.10)}(\zeta)$, that is the debt offers are of $\pm10\%$ of the original debt. The three first sensitivity analysis are with respect to the penalty P, the volatility σ and the growth rate b, see Figure 2.1. We remark that without penalty all the new debt facilities would be accepted, in contrast no debt redemption would occur. The debt redemption area increases with the penalty as can easily forecast. We then have fixed $P=0.08$ for the sensitivity with respect to σ and b, the other sensitivities are obtained for $P=0$. We observe that the dividend threshold increases with P, σ and b.

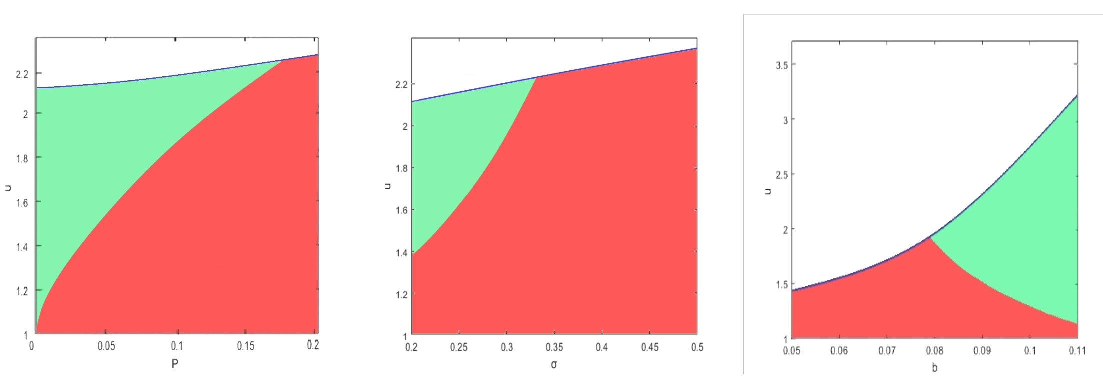

Figure 2.1: Sensitivity of optimal dividend threshold (blue line), debt signature (green) and debt redemption (red) area as function of (a) penalty P, (b) volatility σ and (c) growth rate b.

The most interesting sensitivity is with respect to the intensity λ of the arrival of debt proposal. Figure 2.2 shows that the dividend threshold is a decreasing function of λ. This result could reconcile the model of Jeanblanc and Shiryaev [71] with the indifference law of Modigliani and Miller [92]. In fact, when the intensity λ increases, the optimal threshold $u^\star$ decreases to the level 1, that is the firm has to keep very few cash since all growing opportunities could be financed on the credit markets. On the contrary, when the firm has no access to the credit markets, it needs to keep money to absorb negative shocks, then they pay dividends only if the cash is relatively high, higher than in the case where the credit market is accessible. The same figure in logarithmic scale indicates that the behavior of the convergence of the dividend threshold to 1 is near to a power decay with parameter a little bit smaller than 1.

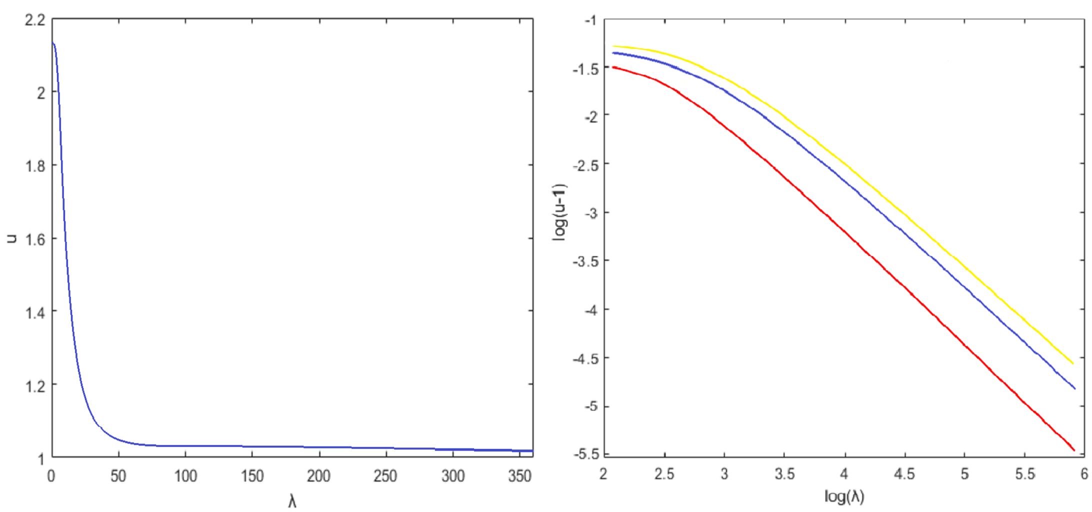

Figure 2.2: Sensitivity of optimal dividend threshold as function of the intensity of the arrival of debt proposal, linear (a) and logarithmic (b) scales. For the logarithmic scale, the optimal dividend threshold is plotted for $\sigma = 0.35$ (yellow line), $\sigma = 0.2886$ (blue line) and $\sigma = 0.2$ (red line).

Finally, we focus on the two regimes case, we assume that $\Lambda_{0,1} = 0.5$ and $\Lambda_{1,0} = 0.2$ in accord with the standard duration of expansion and recession periods, that is five and two years respectively. Figure 2.3 shows the sensitivities of the dividend threshold. When the frequency of the regime changes increases, the dividend thresholds converge toward the same limit. They split as the difference of growth rates increases. Finally, both the dividend thresholds decrease to 1 when the intensity of debt proposal equals infinity, in accord with Modigliani and Miller paradigm.

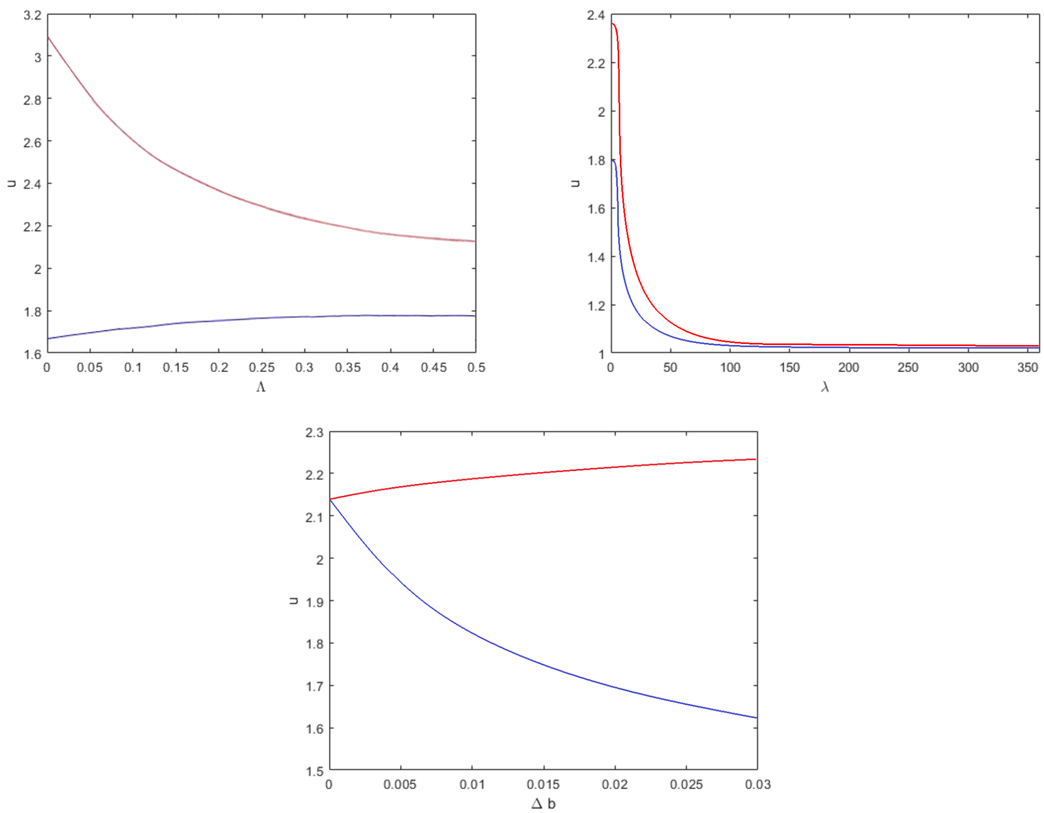

Figure 2.3: Sensitivity of optimal dividend thresholds for regime 1 (red lines) and regime 0 (blue lines) as function of (a) the frequency of regime changes $\Lambda_{i,1-i}$, (b) the intensity of the arrival of debt proposal λ and (c) the difference between the two growth rates b_0 and b_1.

Part II

A lending scheme for systemic risk with probabilistic constraints of failure

A too-big-to-fail is one whose size, complexity, interconnectedness, and critical functions are such that, should the firm go unexpectedly into liquidation, the rest of the financial system and the economy would face severe adverse consequences.

Ben Bernanke (1953-)

Introduction

One of the most relevant changes within the financial world has been caused by the worldwide crisis of 2007-2008. Starting from that breaking event, financial analysts, bank practitioners, applied mathematicians and economists, have been pushed to re-think the models they were used to work with. In particular, it was necessary to stop relying on a series of assumptions turned out to be too far from real markets, as well as from the *new* altered financial worldwide scenario and its changed functioning. Such a *big crunch*, along with its consequences, forced both investors and financial institutions to be aware that almost every financial quantity is exposed to concrete failure risk.

As an example, the standard Black and Scholes (BS) model, whose restrictions on coefficients have been the focus of several studies determining a plethora of alternative and effective approaches, see, e.g., [90, 91], has shown evident limits. In particular, as in the BS model, we base our framework on the geometric Brownian motion. Nevertheless, a major focus of this part of the thesis will be on default probabilities that any financial entity must face.

We underline that such *credit risk* analysis has seen an increasing interest in the theoretical financial community, pushing the development of mathematically rigorous models which take into account both the *risk exposure* factor and related *default events*. Along aforementioned lines, two main approaches have been developed: the totally unexpected failure method, also known as *reduced-form intensity-based model*, see, e.g., [17, ch. 8], and the triggered failure method, also known as *structural model*, see, e.g., [17, sec. 1.4]. Mathematically speaking, the first approach defines the default time as the first jump time of some stochastic process, so that the default event is completely inaccessible to the probabilistic reference filtration modeling the information flow available to traders. After exogenously specifying the conditional probability of default, a typical method to deal with this inaccessibility issue is based on filtration enlargement, see, e.g., [17, 82, 106].

The second approach supposes the default event to be triggered as soon as the value

of the financial entity reaches an endogenous lower threshold. Hence, one of the main issues of the method is the evolution modeling of both the financial entity value and of its capital structure. Therefore, differently from the first mentioned approach, the default time results in being a predictable stopping time with respect to the reference filtration. Let us recall that structural default risk models have been extensively studied in literature, see, e.g., [17, 19, 91, 95]. In what follows we focus our attention on the latter approach, taking into consideration a network of interconnected financial entities, such as banks or general economic agents, who are willing to lend money to each other. We assume that the *bankruptcy event* for a bank occurs when its capital hits a lower barrier whose value is linked to the characterization of the whole system. As a main reference setting we refer to the one introduced in [55], then generalized in [33, 86, 106]. In particular, following [33], we consider a *financial supervisor*, usually referred as *lender of last resort* (LOLR) aiming at guarantee the *wellness* of the financial network, by lending money to those agents who are near to default. At the same time, the LOLR also tries to minimize a given cost function.
Our results also allow to compute the optimal controls for highly complex networks, as the real banking ones. The main novelty of our solution is that, in addition to considering a LOLR who lends money aiming at minimizing a given cost function, we further assume fixed probability constraints the banks have to satisfy at a specific terminal time. From a financial point of view, such constraint implies that the LOLR optimal strategy has to be derived satisfying the assumption that each bank is characterized by a probability of bankruptcy. As in [91], we assume that a bank may fail only at a fixed terminal time, namely it goes under bankruptcy if, at terminal time, its wealth is below a given threshold. This allows us to derive the optimal strategy exploiting techniques related to stochastic target problems. We recall that first results in this direction have been derived in [114], where an *ad hoc* dynamic programming principle has been provided. Later, several papers appeared generalizing such results by considering different constraints schemes, spanning from expectation constraints at fixed time, to almost sure constraints, see, e.g., [22, 29, 30, 31, 63]. In [97], an optimal solution is derived within a similar setting, but without using the stochastic target problem approach. Since in the above mentioned papers examples of concrete solutions are often missing, at the end of this work we consider an example. In particular, we compare our result with the one obtained in [33] limiting, for the sake of clarity, ourselves to a small set of interconnected banks, the case of larger network being of easy derivation. Moreover, because the model construction is strongly based on the mathematical theory of networks, we will exploit its characteristics in order to

derive a *page rank* approach, first introduced in [98], which will be used to determine the *relative importance* of any bank in the network. We then exploit this quantity to decide the admitted probability of each bank's failure, requiring that *important banks* have larger *non-failure probability*, hence adopting a *too big to fail paradigm*.
The part of the thesis is organized as follows: in Chapter 3 we introduce the main setting, giving the mathematical and financial definitions; in particular Section 3.3 we introduce the optimal control problem with probability constraints and we provide its solution; in Chapter 4 we present the Pagerank method for the relative importance of the banks in the network and we apply the derived results to a toy example.

Chapter 3

The general setting

3.1 General framework for systemic risk in financial networks

Let us first introduce the mathematical notation needed to properly treat the general financial scenario we are interested in. In particular, we consider a finite connected financial network identified with a graph $\mathbb{G}$ composed by $n \in \mathbb{N}$ vertices $\mathrm{v}_1, \ldots, \mathrm{v}_n$, corresponding to n banks, and $m \in \mathbb{N}$ edges $e_1, \ldots, e_m$ assumed to be normalized on the interval $[0, 1]$, which represents the interactions between the n banks. In what follows we will use the Greeks letters α, β, $\gamma = 1, \ldots, m$ to denote edges, whereas i, j, $k = 1, \ldots, n$, will denote vertexes. We refer to [38, 39, 95], for further details.

The structure of the graph is based on the *incidence matrix* $\Phi := \Phi^+ - \Phi^-$, where the sum is intended componentwise and $\Phi = (\phi_{i,\alpha})_{n\times m}$, together with the *incoming incidence matrix* $\Phi^+ = \left(\phi^+_{i,\alpha}\right)_{n\times m}$ and the *outgoing incidence matrix* $\Phi^- = \left(\phi^-_{i,\alpha}\right)_{n\times m}$, is defined as follows

$$\phi^+_{i,\alpha} = \begin{cases} 1 & \mathrm{v}_i = e_\alpha(0) \\ 0 & \text{otherwise} \end{cases} , \quad \phi^-_{i,\alpha} = \begin{cases} 1 & \mathrm{v}_i = e_\alpha(1) \\ 0 & \text{otherwise} \end{cases} .$$

In particular, we will say that the edge e_α is *incident* to the vertex v_i if $|\phi_{i,\alpha}| = 1$, so that

$$\Gamma(\mathrm{v}_i) = \{\alpha \in \{1, \ldots, m\} \,:\, |\phi_{i,\alpha}| = 1\} ,$$

represents the set of incident edges to the vertex v_i. We also introduce the *adjacency matrix* $\mathcal{I} = (\iota_{i,j})_{n\times n}$, defined as $\mathcal{I} := \mathcal{I}^+ + \mathcal{I}^-$, where $\mathcal{I}^+ = \left(\iota^+_{i,j}\right)_{n\times n}$, resp. $\mathcal{I}^- = \left(\iota^-_{i,j}\right)_{n\times n}$, is the *incoming adjacency matrix*, resp. *outgoing adjacency matrix*, defined as

$$\iota^+_{i,j} = \begin{cases} 1 & \text{it exists}\, \alpha = 1, \dots, m \,:\, \mathrm{v}_j = e_\alpha(1)\,,\ \mathrm{v}_i = e_\alpha(0)\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

$$\iota^-_{i,j} = \begin{cases} 1 & \text{it exists}\, \alpha = 1, \dots, m \,:\, \mathrm{v}_j = e_\alpha(0)\,,\ \mathrm{v}_i = e_\alpha(1)\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Notice that since $\mathcal{I}^+ = (\mathcal{I}^-)^T$, then we have that $\mathcal{I}$ is symmetric with null entries on the main diagonal.

3.2 Model formulation

Following the financial network setting proposed in [55, 106], and the notion presented in the previous section, we consider a network composed by n nodes, each of them representing a different financial agent, and we denote by $X^i(t)$ the asset value of the i^{th} agent at time $t \in [0, T]$, being $T < \infty$ a fixed positive terminal time. Each node may have nominal liabilities to other nodes directly connected with it. In this case, we denote by $L_{i,j}(t)$ the payment that the bank i owes to the bank j, at time $t \in [0, T]$. Then, we introduce the time-dependent *liabilities matrix* $\mathcal{L}(t) = \left(L_{i,j}(t)\right)_{n\times n}$, defined as

$$\begin{cases} L_{i,j}(t) & \iota^+_{i,j} \neq 0\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{3.2.1}$$

where, as shown in Section 3.1, $\iota^+_{i,j}$ is equal to one if i and j are connected, while it equals zero otherwise. In particular, equation (3.2.1) explicitly states that there cannot be any cash flows between any two banks which are not edge-connected.

At any time $t \in [0, T]$, the i^{th} agent may also have an exogenous cash inflow $F^i(t) \geq 0$. We will denote by $u_i(t)$ the payment made at time $t \in [0, T]$ by the i^{th} bank, whereas $\bar{u}_i(t) = \sum_{j=1}^n L_{i,j}(t)$ is the *total nominal obligation* of node i towards all other nodes. Therefore, if $\bar{u}_i(t) = u_i(t)$, then i has satisfied all its liabilities.

We also introduce the *relative liabilities matrix* $\Pi(t) = (\pi_{i,j}(t))$ defined as

$$\begin{cases} \frac{L_{i,j}(t)}{\bar{u}_i(t)} & \bar{u}_i(t) > 0\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Let us notice that the matrix $\Pi(t)$ is row stochastic, in the sense that $\sum_{j=1}^n \pi_{i,j}(t) = 1$, so that $\pi_{i,j}(t)$ represents the proportion of the total debt at time t that the node i owes to the node j.

Similarly, we can define the cash inflow of the node i as the sum of the exogenous cash inflow $F^i(t)$ plus the total payment that node i receives at time t by other nodes, that is $\sum_{j=1}^n \pi^T_{i,j}(t)\, u_j(t)$, where we denoted the transposed of the relative liabilities matrix and its elements as $\Pi^T = (\pi^T_{i,j}(t))$. We thus have that the value of the i^{th} node at time $t \in [0,T]$ is given by

$$\bar{V}^i(t) = \sum_{j=1}^n \pi^T_{i,j}(t)\, u_j(t) + F^i(t) - \bar{u}_i(t)\,. \tag{3.2.2}$$

Let us now introduce the notion of *clearing vector* as a specification of the payments made by each of the banks in the financial system, see, e.g., [55, Definition 1], [106, Definition 2.6]. In what follows, if not otherwise specified, we will use standard pointwise ordering for vectors in $\mathbb{R}^n$, namely for every x, $y \in \mathbb{R}^n$ it holds $x \leq y$ if and only of $x_i \leq y_i$, for any $i = 1, \dots, n$.

Definition 3.2.1 In the aforementioned financial setting a *clearing vector* is a vector $u^*(t) \in [0, \bar{u}(t)]$ satisfying

- **Limited liabilities:**

$$u^*_i(t) \leq \sum_{j=1}^n \pi^T_{i,j}(t)\, u^*_j(t) + F^i(t)\,, \quad i = 1, \dots, n\,;$$

- **Absolute priority:** that is either obligations are paid in full, or all value of the node is paid to creditors, i.e.

$$u^*_i(t) = \begin{cases} \bar{u}_i(t) & \text{if } \bar{u}_i(t) \leq \sum_{j=1}^n \pi^T_{i,j}(t)\, u^*_j(t) + F^i(t) \\ \sum_{j=1}^n \pi^T_{i,j}(t)\, u^*_j(t) + F^i(t) & \text{otherwise.} \end{cases}$$

Existence and uniqueness of a *clearing vector*, in the sense of Definition 3.2.1, is treated in [55, 106]. In particular, in [55] it is shown that $u^*(t)$ is a clearing vector if and only if

$$u^*(t) = \bar{u}_i(t) \wedge \left(\sum_{j=1}^n \pi^T_{i,j}(t)\, u^*_j(t) + F^i(t) \right)\,. \tag{3.2.3}$$

Equation (3.2.3) can be interpreted as follows: the term $\bar{u}_i(t)$ specifies which $i-$*node* owes to the other nodes at time $t \in [0,T]$, whereas the second term

$$\left(\sum_{j=1}^{n} \left(\pi_{i,j}(t)\right)^{T} u_j^*(t) + F^i(t) \right)$$

represents the cash inflow for the node i at time $t \in [0,T]$. Consequently, *clearing vector* represents the payment at time t of each node: each node pays the minimum between what it has and what it owes. Combining equation (3.2.2) and (3.2.3), we say that the bank i is in *default* if it is not able to meet all of its obligations, therefore the value of a bank equals

$$V^i(t) = \left(\sum_{j=1}^{n} \pi_{i,j}^{T}(t)\, \bar{u}_j(t) + F^i(t) - \bar{u}_i(t) \right)^{+} , \tag{3.2.4}$$

where $(f(x))^+$ denotes the positive part of the function f, so that if $\bar{V}^i(t) \leq 0$, then the bank i is in default, and we set its value to $V^i(t) = 0$.
To simplify the notation, let us define the matrix

$$\widetilde{L} = \left(\widetilde{L}_{i,j}\right)_{n\times n} := L - \operatorname{diag}(u(t)) \,,$$

where $\operatorname{diag}(u(t))$ indicates a $n\times n$ diagonal matrix with the vector $u(t) := (u_1(t), \ldots, u_n(t))$ as its diagonal. The matrix $\widetilde{L}$ has entry $L_{i,j}(t)$ off diagonal, and $-\sum_{j=1}^{n} L_{i,j}(t)$, representing the total payment that the bank i owes at time t to other nodes, on the main diagonal.

Following [86], we assume the liabilities between banks to evolve according the following equation

$$\frac{d}{dt} L_{i,j}(t) = \mu_{ij}\, L_{i,j}(t) \,, \tag{3.2.5}$$

for a fixed positive growth rate $\mu > 0$. We stress that the present setting can be generalized taking L as a geometric Brownian motion. In such scenario the terminal constraint becomes stochastic. Nonetheless, computing the conditional expectation of the terminal constraint, it is possible to recover results analogous the the setting used in the present thesis. We leave this topic to be addressed in a future work. Similarly, we assume the bank i, at any time t, invests the difference between cash inflow ans cash outflow in an exogenous asset $X^i(t)$ whose dynamic is given by

$$\mathrm{d}X^i(t) = X^i(t) \left(\mu^i \,\mathrm{d}t + \sigma_i \,\mathrm{d}W^i(t)\right) \,, \quad i = 1, \ldots, n \,.$$

Moreover, see [86], we introduce continuous (deterministic) default boundaries as follows

$$X^i(t) \le v^i(t)\,, \quad \mathbb{P}\text{–a.s.}\,,$$

with

$$v^i(t) := \begin{cases} R^i \left(\bar{u}_i(t) - \sum_{j=1}^n \pi^T_{i,j}(t)\, \bar{u}_j(t) \right) & t < T\,, \\ \bar{u}_i(t) - \sum_{j=1}^n \pi^T_{i,j}(t)\, \bar{u}_j(t) & t = T\,, \end{cases} \tag{3.2.6}$$

where $R^i \in (0,1)$, $i = 1, \ldots, n$, are suitable constants representing the *recovery rate* of the bank i.

3.3 The stochastic optimal control with probability constraints

In what follows, we introduce the mathematical formulation of our problem, expressing it as an optimal control problem with terminal probability constraint. Furthermore, we provide an analytic solution which allows us to compute the optimal controls.

We consider a complete filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P}\right)$ satisfying usual assumptions, namely right–continuity and saturation by $\mathbb{P}$–null sets. Next results will be applied in Chapter 4 to analyse some financial networks toy models. As in the paper by Capponi et al. [33], we consider a financial supervisor, called *Lender Of Last Resort* (LOLR), connected to any node belonging to the financial network. The LOLR aims at saving the network from default, ant it is assumed to have *full information* about the network state. In particular, at any time t the LOLR can lend money to the bank i, $i = 1, \ldots, n$, so that the controlled evolution of the bank i satisfies

$$\mathrm{d}X^i_\alpha(t) = \left(\mu^i\, X^i_\alpha(t) + \alpha^i(t)\right) \mathrm{d}t + \sigma_i\, X^i_\alpha(t)\, \mathrm{d}W^i(t)\,, \tag{3.3.1}$$

being $\alpha^i(t)$ the loan from the LOLR to the bank i, at time $t \in [0,T]$ and such that $\alpha \in \mathcal{A}$, where $\mathcal{A}$ is the collection of progressively measurable processes $\alpha \in L^2([0,T])$, $\mathbb{P}$–a.s.. In particular, $\alpha(t)$–vector components represent the amounts of money lent to each bank by the LOLR.

In order to derive a closed form solution, we will consider the setting proposed originally by Merton in [91]. Therefore, we assume that default can happen only at some fixed time t_i, $i = 1, \dots, l$, $l < \infty$, hence allowing to only consider constraints defined at terminal time. This allows to avoid introducing strong bonds at each time $t \in [0, T]$.

Let us note that an analogous result can be obtained considering banks allowed to fail at some discrete times $t_1 < t_2 < \cdots < t_M = T$, by separately considering any control problem between two fixed time $[t_i, t_{i+1}]$. This allows to obtain a global control solution by gluing together an ordered sequence of optimal control problems, then exploiting results presented along subsequent sections. Nevertheless, the obtained glued solution is not the optimal one. In fact, in deriving the optimal solution for any time t, one has to consider also possible future evolution of the system. We shall study the latter scenario in a future research exploiting the results here provided, hence deriving the global optimal solution via *backward induction*, as addressed in [36, 100]. Assuming that the LOLR aims at minimizing lend resources implies that he tries to minimize the functional

$$J(\alpha) = \mathbb{E}\left[\frac{1}{2}\sum_{i=1}^{n}\int_0^T \alpha^i(s)^2\,\mathrm{d}s\right]. \tag{3.3.2}$$

Moreover, the LOLR minimizes equation (3.3.2) over the probabilistic constraint

$$\mathbb{P}\left(X^i(T) \geq v^i(T)\right) \geq q^i\,, \quad i = 1, \dots, n\,, \tag{3.3.3}$$

for suitable constants $q^i \in (0,1)$, $i = 1, \dots, n$. For the ease of notation, in what follows, we will drop the index i. Hence, with respect to the agent i, we are going to solve the general control problem, then we apply such result to all banks in the system.

Therefore, let us consider the value of a bank evolving over time according to

$$\begin{aligned} \mathrm{d}X^i(t) &= \left(\mu^i\,X^i(t) + \alpha^i(t)\right)\mathrm{d}t + \sigma^i\,X^i(t)\,\mathrm{d}W^i(t)\,,\\ X^i(0) &= x^i\,, \quad i = 1\,, \dots, n\,, \end{aligned} \tag{3.3.4}$$

and the corresponding default value $v^i := v^i(T)$ at terminal time. Moreover, we will assume that the external supervisor chooses the control α minimizing the following criterion

$$J(t,\alpha) = \mathbb{E}\left[\frac{1}{2}\sum_{i=1}^{n}\int_t^T \alpha_i(s)^2\,\mathrm{d}s \,\middle|\, \mathcal{F}_t\right], \tag{3.3.5}$$

$$s.t. \quad \mathbb{P}\left(X^i(T) \geq v^i \mid \mathcal{F}_t\right) \geq q^i\,, \quad i = 1, \dots n\,. \tag{PC}$$

3.3.1 Reduction to a stochastic target problem

In the current subsection we are going to formally introduce the *Hamilton–Jacobi–Bellman* (HJB) equation associated to the control problem defined in equation (3.3.2), subject to constraint given by equation (3.3.3), hence reducing the related optimal control problem to a stochastic target one. We stress that, due to the structure of the optimal control problem, we will focus on the single agent i. In particular, to avoid heavy notation, if not otherwise stated, we will denote for short $X := X^i$.
Exploiting the value function form given by equation (3.3.5) and by rewriting the terminal probability in equation (PC) as an expectation, namely

$$\mathbb{P}\left(X(T) \geq v \,\middle|\, \mathcal{F}_t\right) = \mathbb{E}\left[\mathbb{1}_{[X(T)\geq v]} \,\middle|\, \mathcal{F}_t\right] ,$$

then we have the following

Lemma 3.3.1 *Given the stochastic optimal control problem with terminal probability constraint* (PC)*, then the terminal probability constraints holds if and only if there exists an adapted sub-martingale* $(P(s))_{s\in[t,T]}$ *such that*

$$P(t) = q\,, \quad P(T) \leq \mathbb{1}_{[X(T)\geq v]}\,.$$

Proof. Let us first prove ($\Leftarrow$): since $P(s)$ is a sub-martingale we have that

$$\mathbb{E}\left[\mathbb{1}_{[X(T)\geq v]}\right] \geq \mathbb{E}\left[P(T) \mid \mathcal{F}_t\right] \geq P(t) = q\,.$$

To prove the converse implication ($\Rightarrow$), let us first denote

$$\begin{aligned} q_0 &:= \mathbb{E}\left[\mathbb{1}_{[X^s(T)\geq v]}\right] , \\ P(s) &:= \mathbb{E}\left[\mathbb{1}_{[X^s(T)\geq v]} \,\middle|\, \mathcal{F}_s\right] - (q_0 - q)\,, \end{aligned}$$

where X^s represents the solution with initial time $s \in [t,T]$, then P is an adapted martingale and the claim follows. □

Let us note that when the probability constraints is *active*, the sub-martingale P is given by

$$P(s) = \mathbb{E}\left[\mathbb{1}_{[X(T)\geq v]} \,\middle|\, \mathcal{F}_s\right] ,$$

hence P is in fact an adapted martingale, and we obtain the new state variable

$$P(s) = q + \int_t^T \alpha_P(s)\,\mathrm{d}W(s)\,, \tag{3.3.6}$$

where α_P, taking values in $\mathbb{R}$, is a new control which, *a priori*, cannot be assumed to be bounded, being derived from the martingale representation theorem.

Remark 3.3.1 Since P represents the probability required to satisfy a terminal constraint, we could have defined P in equation (3.3.6) as

$$P(s) = q + \int_t^T P(s)\,(1 - P(s))\,\alpha_P(s)\,\mathrm{d}W(s)\,,$$

so that P lies in $[0,1]$.

Before explicitly deriving the HJB equation we are interested in, and following [29, 30, 114], let us further simplify our setting by introducing the set

$$D = \{(t,x,q) \in [0,T] \times \mathbb{R}^n \times [0,1] \quad : \quad \mathbb{1}_{[X^i(T) \geq v^i]} - P^i(T) \geq 0 \quad \mathbb{P}\text{–a.s.}\}\,,$$

along with considering the new state variable P, see equation (3.3.6), in such a way that, via the *geometric dynamic programming principle* proved in [114], we can define the value function

$$V(t,x,q) = \inf\left\{\frac{1}{2}\mathbb{E}_t\left[\sum_{i=1}^n \int_t^T \alpha^i(s)^2\,\mathrm{d}s\right] \,:\, \mathbb{1}_{[X^i(T)\geq v^i]} - P^i(T) \geq 0 \quad \mathbb{P}\text{–a.s.}\right\}, \tag{3.3.7}$$

where $\mathbb{E}_t$ is the conditional expectation with respect to the filtration $\mathcal{F}_t$.

Since V is non-decreasing in q, we have

$$V(t,x,0) \leq V(t,x,q) \leq V(t,x,1)\,, \quad q \in (0,1)\,,$$

therefore $V(t,x,0)$ corresponds to the unconstrained problem and its value function is given by $V(t,x,0) = 0$. As regards to the upper bound, we set $V(t,x,1) = \infty$, and we prolong the value function outside $[0,1]$, setting $V(t,x,q) = 0$, resp. $V(t,x,q) = \infty$, for $q < 0$, resp. for $q > 1$.

Let us then introduce the Hamiltonian that must be satisfied by the unconstrained optimal control

$$H^X(x,\alpha,p,Q_x) = (\mu x + \alpha)\cdot\, p + \frac{1}{2}\sigma^2\, x^2\, Q_x + \frac{1}{2}\|\alpha\|^2\,, \tag{3.3.8}$$

where used the following notations

$$\mu x := (\mu^1 x^1,\dots,\mu^n x^n)\,,$$

and

$$\sigma^2\, x^2 := diag((\sigma^1 x^1)^2,\dots,(\sigma^n x^n)^2)\,,$$

being $diag$ the $n\times n$ diagonal matrix.

Intuitively, we are expecting that, when the terminal constraint is satisfied, one can solve the classical associated HJB equation whose Hamiltonian is given in equation (3.3.8), deriving that the optimal control coincides with the unconstrained case. Notice that the optimal solution to the present problem is $\alpha = 0$.

As for the constrained case, and taking into account the new martingale process P, we have to consider the couple

$$\begin{aligned} \mathrm{d}X^i(s) &= \big(\mu^i\, X^i(s) + \alpha^i(s)\big)\,\mathrm{d}s + \sigma^i\, X^i(s)\,\mathrm{d}W^i(s)\,,\\ \mathrm{d}P^i(s) &= \alpha^i_P(s)\,\mathrm{d}W^i(s)\,, \end{aligned}$$

so that we can define the constrained Hamiltonian as

$$\begin{aligned} &H^{(X,P)}(x,\alpha,p,Q_x,\alpha_P,Q_{xq},Q_q)\\ &\qquad = (\mu x+\alpha)p + \frac{1}{2}\sigma^2 x^2 Q + \frac{1}{2}\|\alpha\|^2 + \sigma x Q_{xq}\alpha_P + \frac{1}{2}\alpha_P^2 Q_q\,, \end{aligned} \tag{3.3.9}$$

which should play the role of the Hamiltonian of the associated problem when the constraint is binding. Therefore, the HJB associated to the optimal control reads as follow

$$-\,\partial_t V - \inf_{\alpha\in\mathcal{A}}\inf_{\alpha_P\in\mathbb{R}} H^{(X,P)}(x,\alpha,\partial_x V,\partial_x^2 V,\alpha_P,\partial^2_{xq}V,\partial^2_q V) = 0\,, \tag{3.3.10}$$

where, above and in what follows, for the ease of notation we avoided writing explicitly the dependencies of $V(t,x,q)$.

As mentioned above, α_P could be unbounded, implying that the associated Hamiltonian may be infinite. Since the following holds

$$H^{(X,P)}(x,\alpha,p,Q_x,\alpha_P,Q_{xq},Q_q) \geq H^X(x,\alpha,p,Q_x)\,,$$

to evaluate the minimum of $H^{(X,P)}$ with respect to α_P, we can exploit a first order optimality condition that

$$\alpha_P = -\sigma\, x\, \frac{Q_{xq}}{Q_q}\,,$$

which, when plugged into equation (3.3.9), gives the following minimum for $H^{(X,P)}$

$$\begin{aligned}&\inf_{\alpha_P\in\mathbb{R}} H^{(X,P)} = \bar{H}(x,\alpha,p,Q_x,Q_{xq},Q_q) = \\ &= \begin{cases}(\mu x+\alpha)p + \frac{1}{2}\sigma^2x^2Q_x + \frac{1}{2}\|\alpha\|^2 - \frac{1}{2Q_q}\sigma^2x^2Q_{xq}^2 & Q_q>0\,,\\ (\mu x+\alpha)p + \frac{1}{2}\sigma^2x^2Q_x + \frac{1}{2}\|\alpha\|^2 & Q_q=0\,,\\ -\infty & \text{otherwise}\,.\end{cases}\end{aligned} \tag{3.3.11}$$

It follows that the associated value function introduced in equation (3.3.7) solves the following HJB equation

$$-\partial_t V - \inf_{\alpha\in\mathcal{A}} \bar{H}(x,\alpha,\partial_x V,\partial_x^2 V,\partial_{xq}^2 V,\partial_q^2 V) = 0\,, \tag{3.3.12}$$

subject to the terminal condition

$$V(T,x,q) = \begin{cases}0 & x\geq v\,,\\ \infty & \text{otherwise}\,,\end{cases}$$

where the Hamiltonian $\bar{H}$ is defined as in equation (3.3.11).

3.3.2 The affine control case

In order to obtain a closed form solution for the HJB equation (3.3.11) we will further assume that the admissible controls are of the form

$$\alpha^i(t) = \psi^i(t)X^i(t)\,, \tag{3.3.13}$$

for a fixed constant $\psi^i\in[0,\Psi]$, $\Psi\in\mathbb{R}_+\cup\{\infty\}$. From a financial point of view this implies that the LOLR can decide the interest rate at which the banks assets accrues, allowing the bank to have a higher interest rate to lower the probability of failure.

In the current subsection we derive the explicit solution for the optimal rate ψ

that the LOLR has to give to each bank in order to guarantee their terminal survival probability.

The strategy is as follows: given the structure of the optimal control problem, we can analyse each node i separately, where we ansatz the value function to be of the form $V(t,x,q) = \sum_{i=1}^{n} V^i(t,x^i,q^i)$, where each V^i is regarded as the value function for the optimal problem with respect to the element i. Thus, for each player i we compute the solution to the above problem in terms of contour line of a function $\gamma^i(t,x,q)$, defining first the boundaries of the domain for the value function V^i, then computing explicitly the contour line on the interior of the domain. We stress that the computations will be performed independently for any bank i, nonetheless for the sake of brevity we will omit the index i.

Notice first that given an initial datum and a required survival probability $\bar{q}$, it holds that $\mathbb{P}(X(T) \geq v(T)) \geq \bar{q}$, then the optimal control is given by $\psi \equiv 0$, and the value function $V(t,x,q) \equiv 0$. Therefore we compute three different domains, obtaining, in closed form, two switching curves splitting such domains. The first region Γ_0 is the region in which the constraint is not binding, implying that the optimal control is given by $\psi \equiv 0$. Financially speaking, whenever the value of the bank lies within the region Γ_0, the bank satisfies the LOLR requirement regarding survival probability meaning that it does not need further help to increase its liquidity. Recall that, the more the value of the bank increases, the safer is the bank.

The second region is characterized by the condition Γ_Ψ. In this region the optimal control exceed the maximum rate Ψ that the LOLR is willing to grant, implying that the terminal constraint is not satisfied and the value function V diverges. The last domain, denoted by Γ, is characterized by a binding terminal constraint, and here the optimal control $\psi \in (0,\Psi)$ has to be explicitly computed. Similarly, we will denote by γ_0, resp. γ_Ψ, the switching region between Γ_0 and Γ, resp. between Γ and Γ_Ψ. Regarding Γ let us define the *highest reachable probability* for node i as

$$W^H(t,x) := \sup\{q \,:\, V(t,x,q) < \infty\} = \sup_{\psi \in [0,\Psi]} \mathbb{P}\left(X^{t,x;\psi}(T) \geq v(T)\right) ,$$

where $X^{t,x;\psi}(T)$ denotes the value at time T with initial datum (t,x) and control $\psi \in [0,\Psi]$. It follows that the highest reachable probability is attained when considering the maximum admissible control $\Psi < \infty$, so that by Itô formula and the *Feynman–*

Kac theorem, we have that $W^H(t,x)$ solves the parabolic PDE

$$\begin{cases} W^H(t,x)(T,x) & = \mathbb{1}_{[v(T),\infty)]}(x)\,, \\ -\partial_t W^H(t,x) & = \partial_x W^H(t,x)\,(\mu+\Psi)\,x + \frac{1}{2}\sigma^2 x^2 \partial_x^2 W^H(t,x)\,, \end{cases}$$

whose solution can be explicitly computed to be

$$\begin{aligned} W^H(t,x) &= \mathbb{P}\left(\log\bigl(X^{t,x;\Psi}(T)\bigr) \geq \log\bigl(v(T)\bigr)\right) \\ &= \mathbb{P}\left(W(T-t) \geq \frac{1}{\sigma}\left(\log\left(\frac{v(T)}{x}\right) - \left(\mu+\Psi-\frac{\sigma^2}{2}\right)(T-t)\right)\right) \\ &= \frac{1}{2}\left(1-\mathrm{Erf}\left(\frac{\log\left(\frac{v(T)}{x}\right)-\left(\mu+\Psi-\frac{\sigma^2}{2}\right)(T-t)}{\sqrt{2\sigma^2(T-t)}}\right)\right) \\ &= \frac{1}{2}\left(1-\mathrm{Erf}\bigl(d(\mu,\Psi,\sigma,T-t)\bigr)\right), \end{aligned} \tag{3.3.14}$$

with

$$d(\mu,\Psi,\sigma,T-t) := \frac{\log\left(\frac{v(T)}{x}\right)-\left(\mu+\Psi-\frac{\sigma^2}{2}\right)(T-t)}{\sqrt{2\sigma^2(T-t)}}\,,$$

and Erf denotes the *error function*. For $W^H(t,x) = \bar{q} \in (0,1)$, we have that

$$\frac{1}{2}\left(1-\mathrm{Erf}\bigl(d(\mu,\Psi,\sigma,T-t)\bigr)\right) = \bar{q}\,,$$

and solving for Ψ, we obtain the boundary region in implicit form

$$\Psi = \gamma_\Psi(t,x;\bar{q}) = \left(\frac{\sigma^2}{2}-\mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}}\,, \tag{3.3.15}$$

with

$$\rho := \sqrt{2}\,\mathrm{Erf}^{-1}\left(1-2\bar{q}\right)\,.$$

Thus, for a required probability of success $\bar{q}$, the control problem is not feasible in

$$\Gamma_\Psi = \left\{(t,x)\,:\,\left(\frac{\sigma^2}{2}-\mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}} > \Psi\right\}$$

so that, for starting data (t,x) within the left hand side of $\gamma_\Psi(t,x;\bar{q})$, see equation (3.3.15), the terminal constraint cannot be satisfied, see Figure 3.1. If $\Psi = \infty$, that is the LOLR is willing to give a possibly infinite return rate, any point is controllable, and therefore we can always find an admissible control such that the terminal probability constraint is attained.

As regard Γ_0, computing the no-action region we have

$$\begin{aligned} W^0(t,x) &= \mathbb{P}\left(X^{t,x;\psi_0}(T) \geq v(T)\right) \\ &= \frac{1}{2}\left(1 - \operatorname{Erf}\left(d(\mu,\psi_0,\sigma,T-t)\right)\right), \end{aligned}$$

then by assuming that $W^0(t,x) = \bar{q} \in (0,1)$, we have

$$\frac{1}{2}\left(1 - \operatorname{Erf}\left(d(\mu,\psi_0,\sigma,T-t)\right)\right) = \bar{q}\,,$$

and, solving for ψ_0, we obtain the boundary region

$$0 = \gamma_0(t,x;\bar{q}) = \left(\frac{\sigma^2}{2} - \mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}}\,, \tag{3.3.16}$$

where

$$\rho := \sqrt{2}\operatorname{Erf}^{-1}\left(1 - 2\bar{q}\right),$$

and, as done before, we are left with the following no-action region

$$\Gamma_0 = \left\{(t,x)\,:\, \left(\frac{\sigma^2}{2} - \mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}} < 0\right\}$$

so that, given a starting value $(t,x) \in \Gamma_0$, the terminal constraint is satisfied and the optimal return is given by the null control $\psi \equiv 0$.
At last the action region Γ is the one delimited by Γ_0 and Γ_Ψ, that is

$$\Gamma = \left\{(t,x)\,:0 < \left(\frac{\sigma^2}{2} - \mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}} < \Psi\right\}.$$

Thus, being $(t,x) \in \Gamma$, the controller has to find the optimal control so that the terminal probability constraint holds. By computing the reachability set with fixed constant control $\bar{\psi}$, that is

$$W^{\bar{\psi}}(t,x) = \mathbb{P}\left(X^{t,x;\bar{\psi}}(T) \geq v(T)\right) = \mathbb{E}\left[\mathbb{1}_{[[v(T),\infty)]}\left(X^{t,x;\bar{\psi}}(T)\right)\right],$$

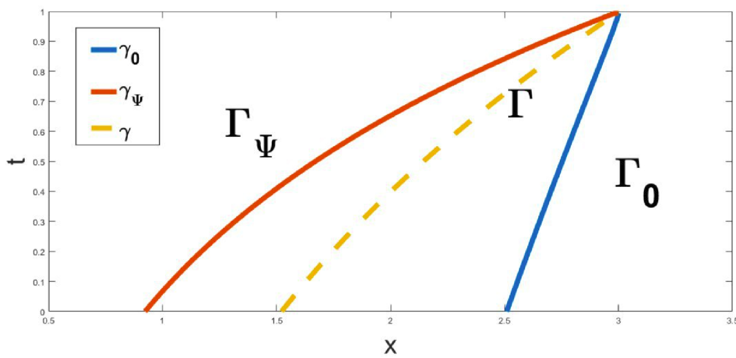

Figure 3.1: Representation of different domains for the optimal control problem.

and proceeding as above, we obtain

$$W^{\bar{\psi}}(t,x) = \mathbb{P}\left(\log\big(X^{t,x;\bar{\psi}}(T)\big) \geq \log\big(v(T)\big)\right) = \frac{1}{2}\left(1 - \mathrm{Erf}\left(d(\mu,\bar{\psi},\sigma,T-t\right)\right) \,, \tag{3.3.17}$$

which implies

$$\bar{\psi} = \gamma_{\bar{\psi}}(t,x;\bar{q}) = \left(\frac{\sigma^2}{2} - \mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}} \,. \tag{3.3.18}$$

What we have obtained so far has to be intended as follows: if the autonomous process $X^{t,x;0}(T)$ already satisfies the terminal probability constraint, then it is optimal to solve the control problem with no terminal constraint, whose solution is given by the null control in the present case.

Therefore, for a fixed $q \in (0,1)$, if $(t,x) \in \Gamma$, the optimal control ψ is given by

$$\gamma_{\psi}(t,x;q) = \left(\frac{\sigma^2}{2} - \mu\right) + \frac{\log\left(\frac{v(T)}{x}\right)}{T-t} - \frac{\sigma\rho}{\sqrt{T-t}} = \psi \,, \tag{3.3.19}$$

see Figure 3.1 for a representation of the above obtained regions. Moreover, along the curve $W^{\bar{\psi}}(t,x)$, the terminal probability of success remains constant, so that the optimal control is given by the constant control $\bar{\psi}$.
Being the optimal control for node i constant along $W^{\bar{\psi}}(t,x)$, then, exploiting equation (3.3.7), the value function for the above control problem reads as follow

$$V^i\left(t,x^i,W^{\bar{\psi}^i}(t,x^i)\right) = (\bar{\psi}^i)^2(x^i)^2\left(\frac{e^{(2(\mu^i+\bar{\psi}^i)+(\sigma^i)^2)(t-T)} - 1}{2(\mu^i+\bar{\psi}^i)+(\sigma^i)^2}\right) \,, \tag{3.3.20}$$

so that we have the following theorem.

Theorem 3.3.2 *The value function for the optimal control problem (3.3.5) is given by*

$$V(t,x,W^{\bar{\psi}}(t,x)) = \sum_{i=1}^{n} V^i(t,x^i,W^{\bar{\psi}^i}(t,x^i))\,, \tag{3.3.21}$$

where

(i) *if* $(t,x^i) \in \Gamma^i$ *and* $q^i \in (0,1)$ *are such that* $\gamma_{\bar{\psi}^i}(t,x^i,q) = \bar{\psi}^i$, *then* $\bar{V}(t,x^i,W^{\bar{\psi}^i}(t,x^i))$ *is given as in equation (3.3.20).*

(ii) *if* $(t,x^i) \in \Gamma_0^i$ *and* $q^i \in (0,1)$, *then it holds* $V^i(t,x^i,q^i) = 0$.

Then V, *as defined above in equation (3.3.21) ,is a classical solution to the HJB equation (3.3.12) on* $\Gamma \cap \Gamma_0$.
Moreover, the optimal control within the class of affine controls is given as in equation (3.3.13), where ψ *is given as in equation (3.3.18)*

Proof. The structure of the optimal control problem gives that the contribute of each node can be treated separately, so that the value function is of the form (3.3.21), where each V^i can be regarded as the value function for the optimal control for the node i alone. As above, for ease of notation, we will omit the index i.

Fixing the node i, it can be trivially shown that for $(t,x) \in \Gamma_0$ we have that $V(t,x,q) = 0$.

Let $(t,x) \in \Gamma$, thus along $W^{\bar{\psi}}(t,x)$, the terminal probability of surviving is fixed, so that explicit computation shows that V, as defined in equation (3.3.20), solves the HJB equation (3.3.12). Observing that the map

$$q \mapsto V(t,x,q)\,,$$

is non–decreasing, together with the fact that $W^{\bar{\psi}}(t,x) > W^{\psi}(t,x)$ for $\bar{\psi} > \psi$, we have that

$$V\left(t,x,W^{\psi}(t,x)\right) = -\infty\,, \quad \psi < \bar{\psi}\,,$$

since the terminal constraint in equation (3.3.7) is not satisfied. Analogously, if $\psi > \bar{\psi}$, then $W^{\bar{\psi}}(t,x) < W^{\psi}(t,x)$. Therefore, as before, the non-decreasing property of V with respect to the third argument q, implies

$$V\left(t,x,W^{\psi}(t,x)\right) > V\left(t,x,W^{\bar{\psi}}(t,x)\right)\,,$$

and the minimum is attained for the control $\bar{\psi}$ implicitly given by equation (3.3.19).

Regarding the value function regularity, notice that it is a classical solution in both region Γ and Γ_0. In order to prove that it is a global classical solution we need to prove that it is regular on γ_0. Let $\bar{x}$ the value on the switching curve γ_0, that is for fixed (t,q), we have that $\gamma_0(t,\bar{x},q) = 0$; then since $\bar{\psi} \to 0$ as $x \to \bar{x}^-$ we have that $\lim_{x\to\bar{x}^-} \partial_x^2 V = 0 = \lim_{x\to\bar{x}^+} \partial_x^2 V$ and $\lim_{x\to\bar{x}^-} \partial_x V = 0 = \lim_{x\to\bar{x}^+} \partial_x V$, hence the value function is differentiable on $\Gamma \cup \Gamma_0$. $\square$

Chapter 4

Application to a network of financial banks

In the present chapter we use the previously obtained results to study a real-world application characterized by an interconnected network of banks. In particular, we will show how optimal solutions previously computed can modify the evolution of such a network. We stress that, for the sake of readability, we will apply our results to a small network, even if, due the fact that the optimal solution is computed in closed form, our results can be easily extended to arbitrary *big* systems.

4.1 PageRank

Before introducing the model, let us introduce an explicit method to address relative importance of a single node in a network. In particular, such an approach will be then used to systematically decide the survival probability for each node.

Let us note that, along previous sections, we have stated an optimal control problem which has been then solved deriving its solution under the assumption that the *accepted probability of failure* q^i is a fixed parameter to be chosen endogenously. In what follows we propose a general, automatic, criterion to deduce the global importance of each node in the system. Next computations exploit results on network analysis already used, e.g., to set the functioning logic of the *Google* research engine, see, e.g., [98]. According to the network formulation introduced in Chapter 3, and using results derived in [98], we show how to score the relative importance of any bank in the network, computing its so called *Page Rank*, allowing us to choose the best survival probability q.

According to the framework described in Chapter 3 let us consider a system of interconnected n banks and related standard *bank enumeration*. Namely, we take into account the usual *one-to-one* correspondence relation between the set of banks and the set of vertexes $V := \{v_1, v_2, \ldots, v_n\}$, referred to as nodes, while $I := \{1, 2, \ldots, n\}$ is the associated set of indexes. Moreover, consider a LOLR strategy in which for each $v_i \in V$ the default probability constraint parameter q^i depends on a predetermined rank R^i associated to the bank$i$, hence representing its systemic importance in the network.

In what follows we are considering graphs as defined in Chapter 3. In particular, to each node $v_i \in V$ corresponds a bank, while to edges connecting nodes $(v_i, v_j) \in V \times V$, we associate the following quantities

$$\gamma^+_{(i,j)} = \frac{c^+ \, L_{i,j} + c^- \, L_{j,i}}{N_j - \min(N) + 1}\,, \quad \gamma^-_{(i,j)} = \frac{c^+ \, L_{j,i} + c^- \, L_{i,j}}{N_i - \min(N) + 1}\,, \tag{4.1.1}$$

where, letting

$$L^+_j = \sum_{i \sim j} L_{ij}\,, \quad L^-_j = \sum_{i \sim j} L_{ji}\,, \tag{4.1.2}$$

$$i \sim j \iff v_i, v_j \text{ are connected,}$$

we define N_j as the net amount of money held by bank j if it would pay its debts at the actual time, i.e. $N_j := X_j + L^+_j - L^-_j$. It is worth empathising that the denominator represents simply a measure for the excess of capitalization for each bank in relation to the less capitalized bank, whose capital is indeed $min(N)$. Moreover c^+ and c^- are two non-negative constants chosen to confer more importance to due debts, resp. to owed credits. For the sake of simplicity, since c^+ and c^- are meant to be weight parameters, we set $c^+ + c^- = 1$. Notice that $\gamma^+_{(i,i)} = \gamma^-_{(i,i)} = 0$ and $\gamma^-_{(i,j)} = \gamma^+_{(j,i)}$, for all $i, j \in I$.

Let us introduce the notion of *outdegree* $\deg^+_\gamma$, resp. *indegree* $\deg^-_\gamma$, for any vertex $v_i \in V$, namely

$$\deg^+_\gamma(v_i) = \sum_{j \in \mathcal{I}} \gamma^+_{(i,j)}\,, \quad \deg^-_\gamma(v_i) = \sum_{j \in \mathcal{I}} \gamma^-_{(i,j)}\,,$$

and normalize the quantities defined in (4.1.1) associated to any couple (i, j) of edges in the graph

$$\overrightarrow{\mathcal{T}}_{(i,j)} = \frac{\gamma^+_{(i,j)}}{\deg^+_\gamma(v_j)}\,, \quad \overleftarrow{\mathcal{T}}_{(i,j)} = \frac{\gamma^-_{(i,j)}}{\deg^-_\gamma(v_j)}\,,$$

corresponding to the ratio of a linear combination on the liabilities between bank i and bank j, and the asset value of bank j. Moreover, we define the matrix $\overrightarrow{\mathcal{T}}$ as the matrix whose entries are $\overrightarrow{\tau}_{(i,j)}$, for $i, j \in \mathcal{I}$, the quantities $\overrightarrow{\tau}_{(i,j)}$ being the weights assigned to each oriented edge.

Therefore, the rating value associated to any node/bank v_i is given by the following recursive formula

$$R_d^i = d \sum_{j \sim i} \overrightarrow{\tau}_{(i,j)} \, R_d^j, \tag{4.1.3}$$

where $d \in (0, 1)$ is a parameter to be chosen, typically $d = 0.85$, see, e.g., [95]. To compute equation (4.1.3), we introduce the so called *Google-matrix*, see, e.g., [95, Ch. 2].

We assume that our network is composed by banks that not only own liabilities if $c^+ = 1$, resp. not only own liabilities if $c^+ = 0$, and at least connected for $c^+ \in (0, 1)$. Of course, banks that are non connected to others in the network, are simply not ranked, since their default cannot affect the system. On the other hand, even if the conditions for $c^+ \in \{0, 1\}$ are not required, they guarantee the boundedness of all the elements of the matrix defined in the next Definition 4.1.1. We stress that, to avoid above restrictions, one can modify the values assigned to edges by equation (4.1.1), e.g., as follows: for $c^+ = 1$ and for every $i \sim j$, define $\widetilde{\gamma}^+_{(i,j)} = L_{i,j}/(N_j - \min(N) + 1) + \epsilon$ as the modified value assigned to the edges.

Definition 4.1.1 (Google-matrix) Let J be a $n \times n$–matrix whose entries are all ones. A *Google-matrix* is a $n \times n$–matrix given by

$$\mathcal{G}_d := \frac{1-d}{n} \, J + d \, \overrightarrow{\mathcal{T}}, \tag{4.1.4}$$

where $d \in (0, 1)$ can be chosen to guarantee irreducibility of $\mathcal{G}_d$, while J is the $n \times n$ matrix whose all entries are 1.

Since the matrix defined in equation (4.1.4) is positive we can apply the Perron-Frobenius Theorem which assures that there exists a *maximum* real eigenvalue $\lambda > 0$ of $\mathcal{G}_d$, indeed λ is the so-called dominant Perron-Frobenius eigenvalue. Moreover, there exists one of the associated eigenvectors, denoted by R_d and usually called *Perron-Frobenius dominant vector*, which is both strictly positive and normalized and whose components represent the rating of each bank. Let us recall that d is usually chosen to be approximately equals to 0.85, see, e.g., [95].

It follows that proposed ranking procedure consists in computing the following series

$$R_d = d \sum_{k=0}^{\infty} (1-d)^k \, (\mathcal{G}_d)^k \, \mathbf{1} \,,$$

where we denoted by $\mathbf{1}$ an n-dimensional vector whose entries are all equal to one.

4.2 A concrete case study

In the present section we consider a systems of banks aiming at computing their ranking. First of all, we are considering a LOLR willing to save banks whose failure would cause insolvency and inability to pay back their liabilities, i.e. $c^+ = 0$, which also implies $c^- = 1$. According to what we have seen along previous sections, see also [95], we fix $d = 0.85$ and we consider a system of banks whose liability matrix and cash vector are as follows, see Figure 4.1 for the associated graph:

$$\mathcal{L} = \begin{bmatrix} 0 & 0 & 10 & 0 \\ 5 & 0 & 5 & 5 \\ 0 & 0 & 0 & 0 \\ 10 & 4 & 0 & 0 \end{bmatrix}, \quad X = \begin{bmatrix} 5.2 \\ 6 \\ 13 \\ 3 \end{bmatrix}.$$

Let us note that proxies for $\mathcal{L}$ and X can be evaluated through the methodology proposed in [50] starting by synthetic data generated by FX markets settlements. However, the aim of this part of the thesis is to focus on the LOLR strategy which should have complete information on the financial market, therefore we do not go into technical details on the estimation procedure.

As explained in Definition 4.1.1, the associated Google-matrix can be easily computed and results as follows

$$\mathcal{G}_d = \begin{bmatrix} 0.0375 & 0.8344 & 0.0375 & 2.3042 \\ 0.0375 & 0.0375 & 0.0375 & 0.9442 \\ 2.9352 & 0.8344 & 0.0375 & 0.0375 \\ 0.0375 & 0.8344 & 0.0375 & 0.0375 \end{bmatrix},$$

where the eigenvalues of the matrix $\mathcal{G}_d$ are $\lambda_1 = 1.2892$, $\lambda_2 = -0.8449$, $\lambda_3 = -0.1472 + 0.3982i$ and $\lambda_4 = \overline{\lambda_3}$. The absolute value of the eigenvector corresponding to the highest eigenvalue is

$$R = v_1 = \begin{bmatrix} 0.3516 & 0.1342 & 0.9177 & 0.1275 \end{bmatrix}^T.$$

The third bank is the one with the highest ranking. Indeed, it is easy to note that its default would cause the default of the first bank and then an insolvency cascade.

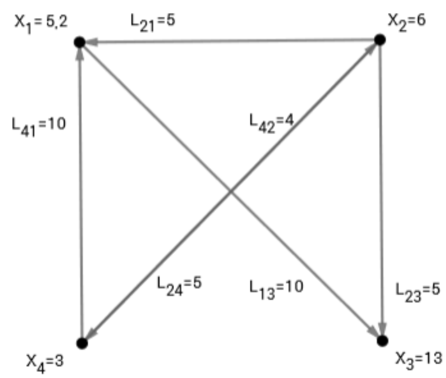

Figure 4.1: Graph representing the system of banks: nodes report the cash value of each bank, while the oriented edges represent the amount of money lend from a bank to another.

Banks (i)	1	2	3	4
X_i	5.2	6	13	3
$\sum_{j\sim i} L_{ji}$	15	4	15	5
R_i	0.3516	0.1342	0.9177	0.1275

Table 4.1: Comparison among the banks rankings.

This is due to the fact that the third bank is systematically more important than the others. Notice that the amount of money due is the most important aspect to be taken into account for the safety of the system. We have reported in Table 4.1 further considerations.

Remark 4.2.1 Looking at Figure 4.1 and Table 4.1, we can see that although the first and third banks are owning the same amount of money to other banks, nonetheless their rankings R are significantly different. This is due to the fact that Bank 3 owns to Bank 1 and its insolvency would probably cause the default of Bank 1. In this example the cascade effect caused by the default of Bank 3 would stop with the default of two banks because of the small dimension of the system, while, on the contrary, such an effect amplifies in big networks.

4.2.1 LOLR strategy under the PageRank approach

In what follows we will describe how to adapt the LOLR problem stated in Section 3.3 to guarantee more flexibility to those banks that are more important for the network's health. Such type of strategies are often referred to as *Systemic importance driven* (SID) strategies, see the next Section 4.3 for more details.

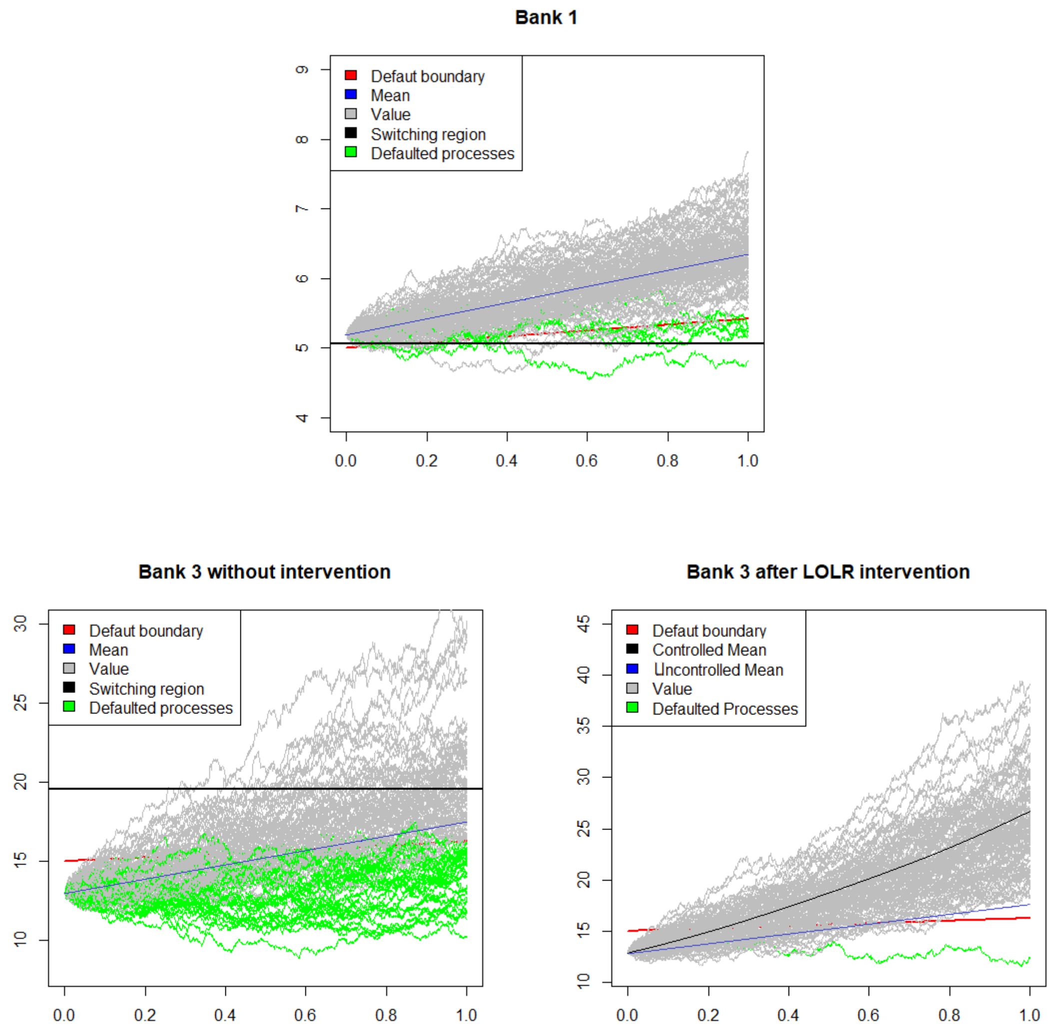

Figure 4.2: 100 simulations for the evolution of Bank 1 (top panel), 100 simulations for the evolution of Bank 3, without LOLR intervention (middle panel) and 100 simulations for the evolution of Bank 3, with LOLR intervention (bottom panel).

We recall that the aim of the LOLR is to minimize the expenditure on banks bailout given by equation (3.3.2) constrained by (3.3.3), i.e. guaranteeing a probability q^i that the bank i will not default. Let us fix an identical probability constraint $q \in [0, 1)$ for all the banks, hence adopting an equality policy analogous to the *max liquidity* (ML) strategy introduced in [33]. We note that a ML strategy guarantees no privileges to any banks, which would lead the LOLR to lend the same amount of money for systematically important banks as for those banks whose failure would not cause a *cascading effect*.

The main idea of the subsequent analysis consists in defining the probability constraints as an increasing function of the rank assigned to each bank. Namely, we have

$$q^i = f(R^i), \quad \text{for } f : \mathbb{R}^+ \to [0, 1) \text{ increasing function} ,$$

where, as seen in Section 4.1, R^i is the ranking of the bank i. Notice that requiring $f' = 0$ the LOLR will again be restricted to the ML strategy.

In [33] was shown that choosing f to be an increasing function leads to a more convenient scenario for the health of networks which have a core-periphery structure, whereas, normally, banks networks have a dense cohesive core, with a periphery less connected.

Coming back to the type of network already defined in Section 4.2, see Figure 4.1, we assume that the LOLR assigns the following probability constraints

$$q^i = 0.9 + 0.05 \cdot \mathbb{1}_{\{R^i > 0.5\}} + 0.04 \cdot \mathbb{1}_{\{R^i > 0.75\}} \, ; \tag{4.2.1}$$

and we perform a one period simulation of the network, see Figure 4.1, taking $t_0 = 0$ and $T = 1$. Let us assume, for the sake of simplicity, that all the liabilities expire at time T, and that they exponentially increase in time with fixed growth rate $r = 0.08$, i.e.

$$L(t) = L\, e^{r\, t}, \quad \text{for } t \in [0, 1] \, .$$

Furthermore, we assume that cash vectors' dynamic evolve according to geometric Brownian motions evolving, namely:

$$\begin{aligned} \mathrm{d}X^1_t &= X^1_t \, (0.2 \, \mathrm{d}t + 0.1 \, \mathrm{d}W_t), \\ \mathrm{d}X^2_t &= X^2_t \, (0.15 \, \mathrm{d}t + 0.25 \, \mathrm{d}W_t), \\ \mathrm{d}X^3_t &= X^3_t \, (0.3 \, \mathrm{d}t + 0.2 \, \mathrm{d}W_t), \\ \mathrm{d}X^4_t &= X^4_t \, (0.05 \, \mathrm{d}t + 0.4 \, \mathrm{d}W_t) \, . \end{aligned}$$

Then, accordingly to equation (3.3.19), we have that the banks' log-switching regions y^i, $i = 1, \dots, 4$, read as follow

$$
\begin{aligned}
&y^1 = 1.622593\,, \quad y^2 = 0\,, \quad y^3 = 2.97332\,, \quad y^4 = 0\,,\\
&q^1 = 0.9\,, \qquad q^2 = 0.9\,, \qquad q^3 = 0.99\,, \qquad q^4 = 0.9\,,
\end{aligned}
$$

recalling that they have to be less than the log initial wealth $X^i(0)$ in order to guarantee the fulfilment of the probability constraint. Therefore, since

$$
\begin{aligned}
\log(X^1(0)) = 1.6487, \quad \log(X^2(0)) = 1.7918\,,\\
\log(X^3(0)) = 2.5649, \quad \log(X^4(0)) = 1.0986\,,
\end{aligned}
$$

we have that the LOLR has to intervene controlling Bank 3. Notice that the LOLR has not to intervene in banks 2 and 4, since they have more credits than debits, hence they cannot face bankruptcy, while the opposite is true for banks 1 and 3. For $q^1 = 0.95$, we would have $\widetilde{y}^1 = 1.6589$ and there would need a LOLR intervention injecting money also in bank 1.

Figure 4.2 (top panel) represents 100 simulation for the evolution of Banks 1 and 3, with and without LOLR intervention. Since the probability of Bank 1 to survive is greater than $q^1 = 0.9$, the LOLR is not going to intervene, whereas indeed its probability to default is approximately 0.062. Clearly, requiring $q^1 = 0.95$ would imply that the LOLR has to intervene lending money to Bank 1. In the middle Figure 4.2, there are represented 100 simulations of the process associated to Bank 3; since $q^3 = 99\%$ and the default probability of Bank 3 is 0.388, the LOLR is going to intervene injecting capital in its cash reserve. After the optimal injection of capital, Bank 3 has probability 0.01 to face the default event, see the lower Figure 4.2, for the representation of 100 simulations of Bank 3 in the case in which the LOLR is going to intervene. Let us underline that the simple case-study we analysed has been set to provide an example as clear as possible, nonetheless, because all the analytical results we derived are in closed form, general complex networks can be theoretically treated as well. Clearly, increasing the graph connection grade, we have an exponential growth in computing the quantities of interest.

4.3 Comparison with the paper by Capponi et al. [33]

As mentioned above, the financial setting has been mainly borrowed by [55] as concerns the lending system formulation, and from [33] for the optimal control problem with an external supervisor aiming at guaranteeing the overall sanity of the system.

This section is devoted to a comparison with [33]. We stress that our assumptions on the optimal control are in the spirit of [33], in the sense that we consider failure at discrete times; also we will not consider a global optimal control, deriving a control for the whole time interval but rather we derive a series optimal control and then gluing together the resulting optimal controls. As mentioned we leave the optimal global control to future research being this latter point mathematically more demanding.

This comparison is significant since their work is based on a similar framework, namely a multi-period controlled system of banks, represented by a network, in which an outside entity, named LOLR, provides liquidity assistance loans to financially unstable banks in order to reduce the level of systemic risk within the whole network of banks. To analyze the systemic risk in interbank networks their work follows a clearing system framework consistent with bankruptcy laws. In particular they generalize the single period clearing system in the paper by Eisenberg and Thomas, see [55], by a multi-period controlled clearing payment system assuming limited liability of equity, priority over equity, and proportional repayments of liabilities after the default event. This generalization leads to a better insight in the propagation and aftershocks of defaults. The main feature in [33] is the comparison between two possible LOLR strategies:

- the *Systemic Importance Driven* (SID) strategy, in which liquidity assistance is available only to banks considered systemically important, i.e. the banks whose default would cause significant losses to the financial system (because of their size, complexity and systemic interconnectedness);
- the *Max-Liquidity* (ML) strategy, in which the regulators aim to maximize the instantaneous total liquidity of the system.

By the analysis of these two different strategies they showed that the SID strategy is preferred when the network has a core-periphery structure, i.e. consisting of a dense cohesive core and a sparse, loosely connected periphery. This is due to the fact that the ML strategy increases the default probability for systematically important banks. Although these two strategies are simplified and do not consider the amount of capital that the LOLR has to inject in the banks network, nonetheless such comparison is useful because the numerical approach fits easily through simulations and systemic risk analysis.

Our work has some important similarities with the one by Capponi et al., in particular we also have considered a finite connected multi-period financial network representing the banks system and the assumptions guaranteeing the consistency with the bankruptcy laws. But, despite this, instead of comparing the two strategies, SID and ML, we considered a LOLR wishing to minimize the square of the lend resources over the probabilistic constraint. Therefore, we did not give an initial budget at disposal to the LOLR as in [33], but took into consideration regulators aiming to find the loan control $\{\alpha^i(t)\}_{i=1,\dots,N,t\in[t_k,t_{k+1}]}$ minimizing the functional given by equation (3.3.2) for each time interval, i.e. $\forall k = 1,\dots,M-1$, ensuring that the probability for each exogenous asset value to be greater than the default boundary is greater than a given constants q^i for each bank $i \in \{1,\dots,N\}$.

Moreover, while [33] is meant to compare two strategies for the LOLR, our approach follows a different path in searching the optimal budget consumption to guarantee a prescribed level of safety of the financial network, given by the parameters q^i $i = 1,\dots,N$. In particular, we do not assume strong constraint over the regulators budget, which depends on the default probability constraint parameters q^i. To switch on a similar comparison as in [33], i.e. considering banks networks of the type *core-periphery and baseline random networks*, and regulator policies of the type *SID and ML*, it suffices to fix the probability constraint depending on the systematic importance of the banks. That is, banks whose failure would cause significant losses to the financial network, because of their size and systemic interconnectedness, should be endorsed with greater default probability parameters q_i. Therefore, our study provides an extension of the admissible policies, through considering an optimal control theory approach.

4.4 Conclusions

In these two chapters of the thesis, we have derived a closed form solution for an optimal control of interbank lending subject to specific terminal probability constraints on the failure of a bank. The obtained result can be applied to a system of interconnected banks, providing a network solution.

We have also shown a simple and direct method to derive the relative importance of every *node* within the studied network. We would like to underline that such a *ranking value* is fundamental in deciding the accepted probability of failure which modifies the final optimal strategy of a financial supervisor aiming at controlling the system to prevent *global crisis* as generalized default.

The results here presented constitute a first step of a wider research program. In particular, in future works we shall consider sequence of *checking times* each of which characterized by possibly different constraints to be considered by the supervisor. In this setting, a solution can be obtained by a backward induction approach, see [36, 100], applied to results here derived. Moreover, as a further development we will consider a framework where the failure can happen continuously in time, hence imposing strict constraints at any time before the terminal one T.

To conclude and anticipate Part III, it is worth noting that central banks and regulators already consider the systematically relevance of important financial institution, and indeed for such institutions and corporations they required different and stricter minimum capital requirements, see e.g. the Enhanced Prudential Standards for US globally systematically important banks of the Comprehensive Capital Analysis and Review (CCAR) of the US Federal Reserve. This technique to tackle the problem concerning the preservation of the health of the financial system is another alternative to the LOLR approach, which also clearly emphasizes the important role that big banks have, but, on the other hand, it forces them to be "too cautious to fails".

Part III

The Default Risk Charge approach to regulatory risk measurement processes

It might be a lot easier to take risks if you're part of a group who will look out for one another.

Andrew Yang (1975-)

Chapter 5

Introduction

The financial crisis obliged the authorities to improve in a dramatic context the banking regulation about the risk management processes and the related capital requirements. The new rules along with the awareness about the weakness of the current practices implied a new exciting research era in the risk management field, covering the whole *end to end* process, from new previously *forgotten* risk sources (systemic risk, liquidity risk, etc) to the attempt to quantify the *model risk*, to the algorithmic effort in order to get faster, granular, reliable risk measures, according to the best reporting standard, such as *auditability*, *drill down* features and so on. In this wide scenario, one of the challenging tasks is to capture the credit risk of the financial instruments, hence removing the classical old boundary where only for the banking book instruments, namely mortgages and loans, the credit risk was measured, while for the financial instruments in the trading book the credit risk calculation was not requested. In fact, it was prescribed only within the so called *Basel 2.5* regulation established starting from 2011, see, e.g., [11], for further details.

The regulation path to this unified view has not yet been completed, but a new relevant step has been developed in the last years with the new Basel framework for the market risk. The new regulation and the new capital charges for the default risk of the trading book portfolio permit to evaluate some alternative computational tools, that can be compared to the usual Monte Carlo approach (MC from now on) used in this area. This part is organized as follows:

- Section 5.1 describes in a self consistent style the financial context and the main objective of the work,
- Section 5.2 introduces the mathematical framework underlying the financial settings,

- Chapter 6, besides containing the outline of the usual model for the default risk calculation, describes the proposal of a quasi exhaustive-heuristic algorithm along with its set-up for real cases.

5.1 The financial context and goal of the work

Since the risk management is a vast field, enriched by a large amount of different applications, spanning from pure practitioners' ones to more theoretically oriented subjects, we have provided some references to avoid any possible misunderstanding concerning both the main goal characterizing the present work and the general context within which it has been considered.

First of all, let us underline that we can distinguish two main different paths in the risk management history. The first one is the scientific risk management, namely the whole set of models, and mathematical techniques developed by both the scientific and the professional community. Such models are often used by banks, according to internal scrutiny procedures mainly aiming at adapting them to real scenarios and contingent decisions.

The second path is constituted by the so called *regulatory risk management*, which is nothing but the set of rules the banks are requested to apply to measure their risks. The latter point is strictly linked to the banks' obligation to have enough capital to prevent them from huge losses. In this direction, a fundamental regulatory framework is the one represented by the Basel Committee on Banking Supervision, or BCBS for brevity. BCBS outlines and updates for each topic the proper framework, then each country adapts them to build their actual regulation, taking care to respect some general not negotiable financial constraints.

The two paths meet and interact, very frequently. Typically, new techniques and models are accepted in the general regulation framework, provided there exists a robust awareness about them.

A milestone example is represented by the competition between the Value at Risk (VaR) and Expected Shortfall (ES). The VaR was adopted as the official market risk measure in 1996 by BCBS, see, e.g., [9]. Then, the academic community pointed out its drawbacks, such as the lack of sub-additivity property, see, e.g., [2, 3]. Nevertheless, such risk measure has been adopted so far and is going to be revised within the incoming *Fundamental Review of Trading Book* (FRTB) framework forecasted to enter in force from 2022, when the VaR will be replaced by the ES, see [13] for further details. In this scenario, our work has been mainly focused to consider the credit risk in the trading book, namely to consider the *risk* associated to a portfolio of financial instruments held by the banks, such as bonds and equities. Even if the credit risk field is a very large area, it is possible to split it in two main research areas. The first one is represented by the development and analysis of suitable *default models* for each counterparty, along with the estimation of related default probabilities. The second major theme is the so called *portfolio credit risk* which, roughly speaking, aims to optimize the way to collect each debtor's risk to obtain a consistent risk measure for the whole portfolio. The latter implies a difficult point to be solved, namely how to properly infer the default correlations among different debtors that have not yet been observed at the calculation date. Let us recall that, concerning the default probabilities evaluation, the most celebrated contributions are the structural models by Merton, see [89], while, to what concerns the portfolio credit risk, a fundamental seminal work dates back to the Vasiceck contributions, see [115]. The banking industry has tried to exploit these first scientific contributions, working on their mathematical peculiarities, to develop more effective calculation processes, as in the case of the *binomial-based model* by *Credit Suisse* and the *Credit Metrics model*, for further details see [96]. The embedding of previously mentioned developments within the *Credit Risk* general framework started in in 2006, when the Basel 2 regulation admitted the statistical models to measure the credit risk of the portfolio. Until that date, only the standard models where allowed, which implied to consider a set of grids of coefficients, each of which applied to a different exposure category. Along this regulatory line, the statistical model prescribed by the BCBS has been the Gordy one (2003), also known as the *Asymptotic Single Risk Factor* model (ASRF). The ASRF model is characterized as follows: each debtor (i) has a behavior Y_i defined by

$$Y_i = b_i \cdot X + \sqrt{q - b_i^2} \cdot w_i \,,$$

where $w_i, X \sim N(0,1)$, with $\rho(X, w_i) = 0$. In particular, X is the single systematic risk factor and all the debtors depend on it by the *factor loading* b, while the specific features are summarized in the independent Brownian-type noise represented by w_i. The counterparty defaults if Y_i is below a threshold K, given by $K = \Phi^{-1}(P)$, P being the default probability of (i) that is estimated by other statistical models. We underline that ASRF model is very appealing for the regulator, in fact, if the credit portfolio has many positions with exposures amount E_i and it holds the perfect granularity property, i.e.

$$\frac{E_i}{\sum_{j=0}^{n} E_j} \to 0 \quad \forall i \,,$$

then the VaR of the whole portfolio can be analytically obtained by summing up the risk contributions coming from each position. It is worth to mention that, in real markets, the perfect granularity property does not hold perfectly, nevertheless if the credit portfolio is very large and not much concentrated, then the analytical VaR formula represents a good approximation. We have to consider that typical banks' positions belong to two broad categories, namely: the *banking book* and the *trading book*. While the banking book mostly consists of classical credit products, such as loans, mortgages and so on, the trading one is mainly constituted by structured financial instruments, such, e.g., bonds, equities, derivatives. For the sake of simplicity, we do not analyze the accounting perspective that allows to classify also some financial portfolio in the banking book category, if there is not a trading purpose for those positions. One of the most important weaknesses of the Basel 2 regulation lies exactly in the *banking vs trading book classification* challenge. More explicitly, the default risk has to be measured within the regulatory framework only for the banking book portfolio, not for the trading book. As an example we consider a portfolio of plain vanilla bond. Only the interest rate risk, which is referred as the *generic risk* in the BCBS language, and the spread risk, or *specific risk* from the regulation point of view, are captured by a 10 day 99% VaR, without any measurement and capital constraint assignment for the portfolio default risk.

From the 2007-2008 crisis we learned that also the big banking institutions and large corporates can fail, implying a huge amount of losses in the trading book portfolio, also because of possible *contagion phenomena*, see, e.g., [15] and Part II. These types of losses can not be absorbed by the bank capital, at least if no provision for the risk has been previously stated. To solve this gap an updated regulatory directive, namely the *Basel 2.5* one, has been developed, see [12] and [11] for further details, allowing, in particular, for a new risk parameter called *Incremental Risk Charge* (IRC).

This new risk measure is a 99.9% VaR with 1 year horizon. It takes into account both default risk and migration risk, or *down grade*, and it has been mainly prescribed for bonds-type instruments. The Basel 2.5 regulation came into force in 2012, accompanied by a lot of criticisms because of its overreaction properties that obliged banks to immobilize huge amount of their capital. In particular, the IRC measure is an example of *risks double counting*. In fact, the migration risk captured by the 99.9% 1 year VaR clearly overlaps the *old* classical risk measure for the market risk, i.e. the 99% 10 days VaR. Such an issue generates too high risk figures, with banks unable to efficiently allocate the requested capital. Therefore, BCBS published an updated version of the previous regulatory rules, providing the *fundamental review of* *the trading book* (FRTB), see, e.g., [13], which, even if becoming into effect just starting from 2022, has immediately pushed banks to be compliant with it, by applying its the new rules in advance within their reporting and limits system as well as to try to develop properly adapted internal models. Therefore, the FRTB has allowed to overcome some of the previously present *financial misunderstandings*. The IRC has been replaced by the *Default Risk Charge* (DRC, formally Incremental Default Risk charge or IDR).

The main features of DRC are:

- It is still a 99.9% 1 year VaR.
- Only default risk is considered, while migration risk is removed.
- Equity style positions have to be considered.
- The default model must be a 2 *systematic factors model*, to overcome the drawbacks of ASRF. In fact, since trading book often consists of a relative small number of concentrated positions, namely just dozens or hundreds of debtors, instead of set many thousands, the ASRF proxy could provide too rough estimates.
- The default correlations must be jointly estimated, hence taking into consideration both equity prices and spread movements.

An excellent review of the FRTB regulation can be found in [83], while the new DRC challenges are clearly explained in [102]. Within this scenario, we do not aim to suggest what systematic risk factors have to be selected, how to estimate the default correlations, how to map illiquid instruments, or what is the proper proxy to be considered. Instead, in the global DRC calculation work-flow, we try to innovate in the final quantile calculation procedure. In other words, once the model has been stated, the positions have been classified and we know the joint default probabilities of all the debtors, we want to compare the classical Montecarlo approach with deterministic exhaustive or near exhaustive procedures.

It is worth to observe that the more the quantile level is extreme, 99.9% for DRC, the more the Montecarlo empirical quantile can suffer of high variance estimation error, as outlined, e.g., in [43].

Then a huge number, such as 10 or 100 millions, of heavy simulations are run by the banks to get convergence of the empirical estimator, implying a high time consuming procedure which is also not so easy to set-up.

At the best of our knowledge, our analysis innovates the already present literature on the subject, showing a new proposal that we are confident can be a very promising alternative to the existing ones.

5.2 The mathematical setting - Model for the IRC and DRC measure

To focus on the statistical and algorithmic problem, we skip some of the several technical details of the regulation, and we simplify a bit the complex model. Roughly speaking, most of the model for IRC are *structural* models, where the default event of each issuer is related to some *background* risk factors. Moving from *IRC* to *DRC*, the model for the portfolio losses could remain the same, only the events to be considered are changing.

For a general overview of the structural model *á la* Merton, see the seminal paper in [89]. For the practical implementation in the industry, the benchmark model is the *CreditMetrics* model, see [96].

For the sake of simplicity we focus now on the most common model in the banking practice, avoiding too many theoretical definitions and preliminaries.

Notation

- J: the number of issuers (of bonds, equities, ...) in the trading portfolio. Usually for medium banks J could be of some dozens, more than one hundred only for very large banks. Moreover, we have a *concentration* effect, e.g. with the top 10 issuers one has a relevant fraction of the whole portfolio value.

- MtM_j: the *mark to market*, or present value, of the instruments issued by the *j-th* issuer, then $MtM_j = \sum_{i=1}^{I(j)} MtM_{i,j}$, where $I(j)$ is the number of instruments hold in the portfolio and issued by issuer j.

- DP_j: the default probability of the *j-th* issuer.

- R_j: the rating level of the *j-th* issuer at the evaluation time. Usually (it is merely a convention) the lower the rating code, lower the default probability. Usually the ratings are useful tools to group the default probability levels, hence one can write $DP(R_j)$ instead of DP_j to make explicit this mapping process.

- rr_j: the *recovery rate* for the *j-th* issuer once the default event happens. This value is a fraction in the range $[0, 1]$, but usually it is quite close to 35%, 40% for bond instruments, 0% for equity instruments. In the practice, we can not have an estimated recovery rate for each issuer, before it defaults, then the recovery rates are grouped by historical data and clustering all the default events by sector (financial, corporates, govies) and / or geography. We indicate with $s(j)$ the sector to which the issuer j belongs, we can write more explicitly $rr(s_j)$. The complement to 1 of the recovery rate is the loss given default (LGD in short), $lgd_j = 1 - rr_j$.

- L_j: the loss due to the default of the *j-th* issuer. Here we do not focus on the joint stochastic dynamics of market (interest rate, forex exchange, equity prices) and credit (default, spread) risk factors, hence we simply use the expression $L_j = MtM_j \cdot lgd_j \cdot F_t$, where F is the forward factor over a time horizon t and is equal to 1 at the default time is the MtM is supposed not to change significantly. Hence here L is not a random variable, but the loss once the default occurs. Furthermore, as usual from the *practitioner's point of view* as well as within the Basel regulation models, we replace the random loss given default fraction with its expected value LGD. Therefore, L can be considered as an expected loss, conditioned to the default event.

- RL: number of rating levels.

- $PL_j\,(rl)$ is the profit (or loss) for the holdings in the *j-th* issuer if its rating moves from the current level to the level $rl = 1, 2, \ldots, RL$. Practically, a further mapping process is performed, and for each combination (s, rl_1, rl_2) of rating migration for a given sector s a spread movement is established, let be $\Delta spread_{\{s,rl_1,rl_2\}}$. We stress the fact that, since the lgd_j do not differ significantly from a j to another, the vector of the losses L_j may be viewed as almost proportional to the vector of the mark to markets MtM_j.

With the above equipment, we can finally write the portfolio loss due to the default events, let $Loss_P$

$$Loss_P = \sum L_j \cdot \mathbb{1}_{D(j)}\,,$$

being $\mathbb{1}_{D(j)}$ the *indicator* function of the default event. Despite the very compact expression, the calculation process of the risk figures (quantile, expected shortfall, etc.) are very involved, because of the complex parameters estimation process one has to set and to maintaining, and for the dependency structure between the issuers defaults. The *IDR* measure is defined as $VaR\left(Loss_P, 99.9\%, 1Y\right)$.

If we are interested to the broad profits and losses profile that we could observe in the portfolio because of the migration events we have

$$PL_P = \sum PL_j\left(R_j\right).$$

In the above expression most of the complexity is in the chain of possible migration events, the associated PL, the dependency between the different issuers migration.

The above quantities are the key point of our application in the next section. At the end, the random variable PL_P is a discrete one. How many possible outcomes could it have? Given the J issuers, and given the number RL of rating levels, we have that the cardinality of the outcomes of the random variable PL_P equals RL^J, i.e. $\#\{\text{outcomes}(PL_P)\} = RL^J$.

Here we are analyzing the issue by a strict *cardinality* perspective of the points in the space Ω where the elementary events ω_i take place. Of course we could have different events with the same numerical value of PL_P or some events that have very negligible probabilities. Anyway, with practical cases parameters it comes out a number of outcomes that can not be dealt satisfactorily in an exhaustive fashion.

Here *exhaustive* means that we could theoretically calculate the exact distribution of the random variable PL_p, i.e. its outcomes $\{x_i\}$ and the related probability masses $\{q_i\}$, and then to obtain the quantile by properly cutting the cumulative (discrete) distribution function at a given level.

With $J = 20$, $RL = 10$ we have 10^{20} possible outcomes. Obviously this cardinality cannot be actually managed due to 3 main reasons:

1. *computational*, i.e. to calculate all the $\{q_i\}$ probability masses;
2. *ordering*, i.e. to order the $PL_P(i)$ to get the quantile;
3. *storage*, i.e. to write in a database the whole input-output combinations for auditability purposes.

But thanks to the IDR new regularization the number of outcomes decreases dramatically, hence one can wonder if some exact calculation can be performed. Let us recall that in this case we have "only" 2^J elementary outcomes. With the above parameters $2^{20} \simeq 1,000,000$. Then at least theoretically we can try to face it by exploiting its feasibility, computational time and so on.

Chapter 6

Evaluation approaches

In this chapter we will face our aim to determine the value of the minimum capital required to stand up to losses due to the 99.9% worst possible scenarios of issuers defaults with shares in the portfolio considered.

This chapter will be divided in two main sections that correspond to the description of a numeric algorithm for a relatively small portfolio dimension and the description of a simulation algorithm for a portfolio with larger dimension. Because of the heuristic approach for the larger dimension portfolio, the second section will be integrated by a statistical comparison.

For both cases it will be considered a portfolio consisting of J issuers; J will be the dimension of the problem. Practical experience has shown that common portfolios usually satisfy some characteristics on the loss vector $L = [L_1, L_2, \ldots, L_J]$ and on the default probability vector $DP = [DP_1, DP_2, \ldots, DP_J]$ which are summarized by the following properties:

- We have a usual *concentration effect* in the asset allocation, that is more in the govs bonds and less in the corporate bonds. Furthermore we can heuristically assume that the loss determined by the default of the 10% of the issuers among the larger ones, should be at least the 90% of the maximum possible loss in the portfolio, $\sum_{j=1}^{J} L_j$. Here the values 10% and 90% are an example, and other similar assumptions on the concentration may be made. As we outlined in section 3, we have that $L \approx c \cdot MtM$, i.e. the 2 vectors have approximately the same distribution. Hence from now on we simply use $L = (L_j)$;

- L can be satisfactorily approximated by a *Beta distribution*. This approximation is justified since the global loss $L = \sum L_i$ is bounded, which is coherent with the Beta distribution characteristics and, moreover, by updating its parameters, the distribution behavior can be adapted to fit different loss shapes. It is also worth to mention that the Beta distribution is rather popular within financial applications to describe the loss given default quantity, which implies that it is a widely accepted random model in the credit risk management practice. We would also like to underline that the loss is the sum of weighted, by the portfolio fractions, loss given defaults. From this point of view, one can also consider different alternative probability distributions, as the Gamma or Log-normal ones, nevertheless they are more suited for unbounded losses cases, as, e.g., when considering the operational risk field, as pointed out in [94].

- Issuers with greater L_j have a lower default probability DP_j, because of the prudent asset allocation of the bank's portfolio.

- Default probabilities are in the range $[0, 10\%]$.

It is worth to mention that such empirical facts admit some exceptions, nevertheless the above points have some rather intuitive rationale. Typically, if we look at the bond portfolio of a typical European commercial bank, it is characterized by long positions with a relevant part of government bonds, and, in this category, the first holding is for the national country bonds with *good tradition*, e.g., *BTP* for Italy, *bonos* for Spain, *bunds* for Germany, etc. Moreover, all banks have some limits in the asset allocation policies, where some strict upper bounds are assigned for instruments and issuers with low rating, such as B, BB, etc. Furthermore, some limits are given also by sector, where corporate bonds, typically with higher default probability, cannot exceed a given threshold. Hence we have the double concentration effect mentioned above. As an example, we exhibit the portfolio composition, as of 2016, December, of one of the largest Italian banks. For the sake of clarity, here the weights are relative weights, and the total value of portfolio is about 20 billions €:

- Numbers of issuers in portfolio: about 90;

- Weight of the first bond issuer, or *Italian Republic* issuer: 84%;

- Weight of the first 5 holdings: 91%.

Figure 6.1 represents the plot of the cumulative weight of the bond issuers. Due to the huge variety of banks and issuers in the market, there is not a comprehensive empirical analysis in the literature, but several specific studies, most of which have been conducted by central banks and financial authorities that have access to data

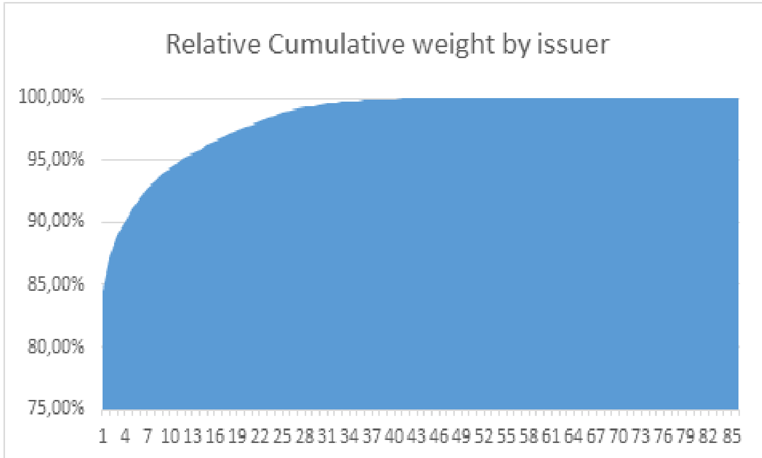

Figure 6.1: Plot of the cumulative weight of the bond issuers.

where generally there is not a granular disclosure, namely the bank portfolios holdings. A very recent deep analysis can be found in [87].

These portfolio management techniques explain both the concentration effect and the negative relationships between the PDs values and the exposures sizes. To this extent, let us recall that the market spread values for bonds can be misleading in the perception of the default probability, namely for the government bonds. In fact the FRTB regulation strictly prescribes to not use the market, risk neutral, PDs implied by the spread, but instead the statistical real world estimation, see, in [13], the 352 BCBS paper, par.186.(s). Therefore, the high spread levels and volatilities characterizing some EU countries, where the banks have the most relevant holdings, do not imply high PDs in the DRC calculation, namely: the real probabilities are very low, hence matching the negative relationship principle.

To obtain a vector L satisfying these properties, for each single issuer $j = 1, \dots, J$ we simulate the values of L_j through a beta distribution $f(\alpha, \beta)$, with parameters $\alpha = 1/15$ and $\beta = 5$, of the default probability, DP_j, that will be proportional to its L_j. In Figure 6.2 we split the range for the loss values into 40 sub-segments and represented the frequencies of occurrence of 1000 losses L_j, simulated by a $\text{Beta}\big(5, \frac{1}{15}\big)$, and we did the same for the default probabilities DP_j, which are inversely proportional to the value assumed by the correspondent L_j. Figure 6.3 represents the scatter plot of $J = 200$ points of a single portfolio simulation for the couple (DP, L) to show the tendency of an inverse relation between the amount of losses and the default probabilities.

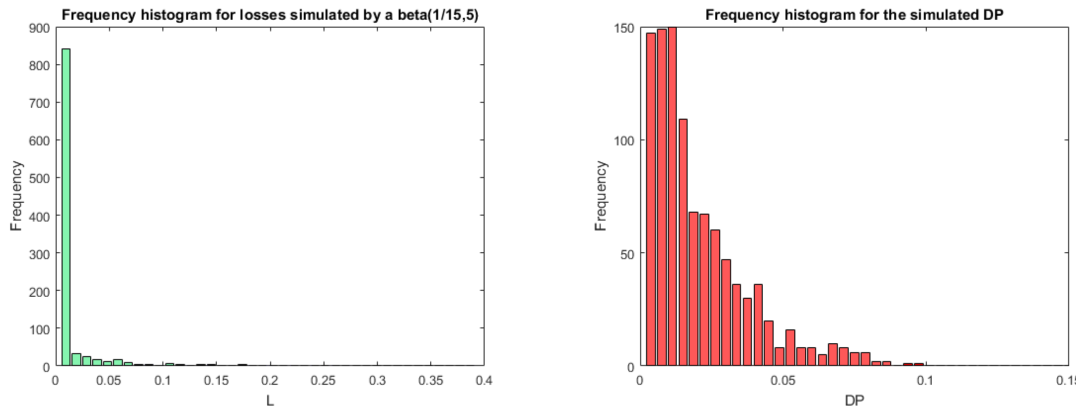

Figure 6.2: Histogram representing the frequency of occurrences of 1000 values of L and DP. We remark the fact that banks have different portfolios from each other, and that the portfolios are time changing, therefore taking into consideration simulated portfolios is not less accurate than considering samplings that are relatively small with respect to the wide empiric variety.

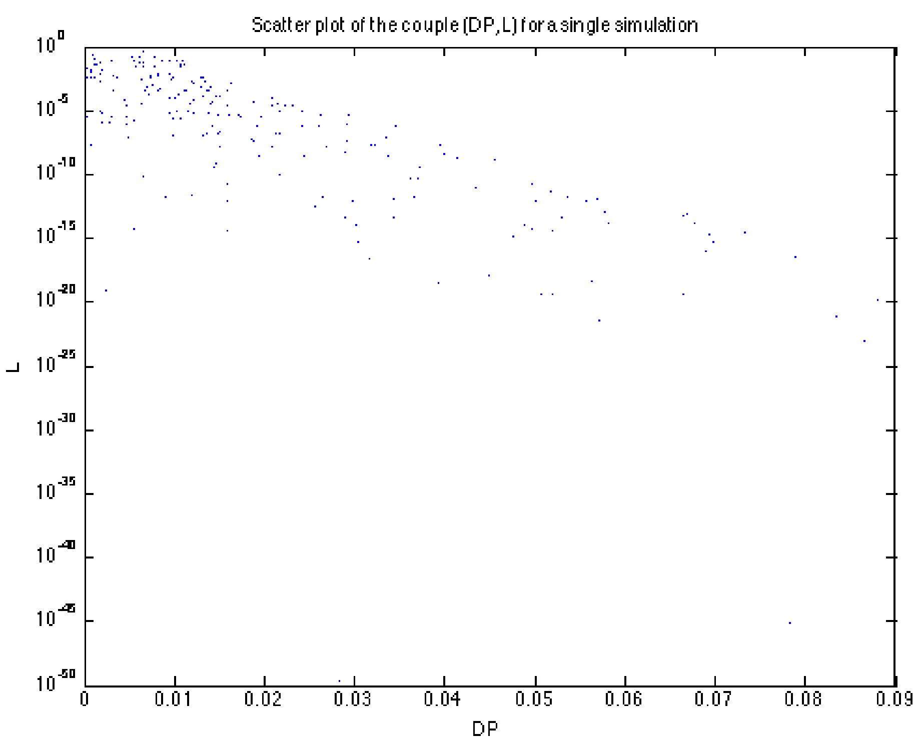

Figure 6.3: Scatter plot of a single simulation of $J = 200$ points of the couples (DP, L) with logarithmic scale for the y-axis.

6.1 Determination of the minimal capital required through comprehensive exact approach

Let us start from a simple case. We consider a portfolio P consisting in shares of $J = 20$ issuers, assuming that the loss vector L and the default probability vector DP are such that P satisfies the properties derived by practical experience.

Since every issuer is associated to a DP_j, we can compute the conditional default probabilities determined by defaults of combinations of issuers. In this part of the thesis we will assume that the issuers are uncorrelated, therefore it is the starting point for further analysis. The reason for this independence assumption on the issuers is mainly motivated by the fact that this study is meant to emphasize the calculation procedures and the machine execution effort, instead of the statistical model that binds the issuers.

Having $J = 20$ issuers in the considered portfolio, implies that the number of possible conditional default scenarios is

$$N = 2^J = 1048576,$$

and if we have n issuers that will default, the number of possible combinations is given by the binomial coefficient $\binom{J}{n}$. To each default scenario we can associate its conditional default probability CDP. For n issuers that will default, we consider the index set I with dimension n such that each element in I represents the index of a issuer that will default in the simulated scenario. Then the indexes for the solvent issuers belong to the set

$$\{1, 2, 3, ..., J\} \setminus I.$$

The Conditional Default Probability associated to the index set I is

$$CDP(I) = \prod_{i \in I} DP(i) \times \prod_{\substack{1 \le j \le J \\ j \notin I}} \big(1 - DP(j)\big), \tag{6.1.1}$$

and its corresponding loss is

$$Loss_P(I) = \sum_{i \in I} L(i)\,. \tag{6.1.2}$$

We want to determine the 99.9% worst case, therefore we sort all the possible $Loss_P$ in descending order as elements of the vector $Loss_P^{\text{Dec}}$, and the CDP with the same order as the sorted $Loss_P(I)$ as elements of the vector $CDP^{Loss_P\text{-sort}}$. Hence the first element of the vector for the reordered $Loss_P^{\text{Dec}}$ is the loss determined by the default of every issuer, i.e.

$$\begin{aligned} Loss_P^{\text{Dec}}(1) &= Loss_P\big(\{1, 2, \ldots, J\}\big) \\ &= \sum_{i=1}^{J} L(i)\,, \end{aligned}$$

and the first element of the $CDP^{Loss_P\text{-sort}}$ reordered as the $Loss_P^{\text{Dec}}$ is

$$CDP^{Loss_P\text{-sort}}(1) = \prod_{i \in 1}^{J} DP(i)\,.$$

To provide an example, let us assume that the issuer with lower L is the j-th one, then the second element for the sorted $Loss_P^{\text{Dec}}$ would be the summation of all the issuers L except the one of issuer j, and in general the second element for the reordered $CDP^{Loss_P\text{-sort}}$ is

$$CDP^{Loss_P\text{-sort}}(2) = \prod_{\substack{1 \le i \le J \\ i \neq \text{argmin}_j L(j)}}^{J} DP(i) \times \big(1 - DP(\text{argmin}_j L(j))\big)\,,$$

and so on for all the 2^J possible scenarios.

The IDR we are searching for is the element of $Loss_P^{\text{Dec}}$ corresponding to $\hat{k} - 1$, where $\hat{k}$ is the lower index such that the cumulative probability satisfies

$$\sum_{i=1}^{\hat{k}} CDP^{Loss_P\text{-sort}}(i) > 0.1\%\,.$$

Since *CDP* is a discrete variable, we cannot consider the usual definition of quantile of a continue random variable, instead we read the 99.9% quantile of the discrete variable *CDP* as the higher index $\hat{k}$ such that

$$\mathbb{P}(Loss \le Loss_P^{\text{Dec}}(\hat{k} - 1)) \le 99.9\%\,,$$

where $Loss : \Omega \to \mathbb{R}^+$ is a discrete random variable which takes values accordingly to (6.1.2) for all the possible combinations in $\Omega = \{0, 1\}^J$, where one and zero entries represent default and non-default for each issuer, e.g. an outcome $\omega \in \Omega =$

$(0, 0, 0, 1, 0, 0, 1)$ would represent the default of the fourth and seventh issuer in a $J = 7$ dimensional portfolio. Therefore for simplicity of notation we consider this approximation as the desired percentile, i.e.

$$\mathrm{IDR}(99.9\%) = Loss_P^{\mathrm{Dec}}(\hat{k} - 1)\,. \tag{6.1.3}$$

To conclude the analysis of the deterministic approach, let us underline that the exact algorithm we have presented, works satisfactorily for low dimension portfolios and uncorrelated issuers. While, if the dimension J representing the number of issuers starts to be greater than 30, then the computational efforts needed to evaluate the IDR for the 99.9% starts to be rather demanding, since the algorithm complexity grows exponentially in J, i.e. $T(J) = O(2^J)$.

6.2 Determination of the minimal capital required through heuristic technique HR

As we have seen in the previous section, increasing the portfolio dimension the execution time for the exact algorithm becomes prohibitive. For this reason, for higher dimensions we simplify the complexity of the IDR evaluation problem through an estimate from a regression model. Let us denote by J^* such higher dimensions in order to mark the difference between this estimation problem and the previous one. We also denote as P^* the portfolio consisting in shares of J^* issuers, and will be referred as *Portfolio Target*.

Considering n simulated J dimensional portfolios, we estimate the relationship between y, ratio of IDR(99.9%) with respect to the maximum possible $Loss_P$ given by default scenarios, i.e.

$$y = \frac{\mathrm{IDR}(99.9\%)}{\sum_{i=1}^{J} L(i)}\,, \tag{6.2.1}$$

and two concentration indexes. We remark the fact that y is a relative IDR which is a dimensionless index. Indeed, the basic idea behind the evaluation of y is to start estimating the IDR of the Portfolio Target P^* through the concentration indexes and then switching to a dimensionless scale with respect to its dimension J^*. To be more precise, supposing that the quantile is a well estimable fraction, given by the concentration curves of L and DP, we extrapolate the behavior observed in the lower dimension $J < J^*$ to adapt it to the Target problem in dimension J^*. We consider the concentration indexes that measure how the total $Loss_P$ and the sum of all the default probabilities are divided through the issuers. Let us explain how we get these concentration indexes. First of all we sort the issuers by their losses L. Then, since to each issuer is associated a L and a DP, from the reordering we obtain two vectors:

1. L^{Inc}, which has as elements the L arranged increasingly,
2. $DP^{L\text{-sort}}$, which has as elements the DP corresponding to the losses L with same index in the vector L^{Inc}.

Let us denote by $Q_1^{(p)}$ and $Q_2^{(p)}$ the concentration indexes for L^{Inc} and $DP^{L\text{-sort}}$ given by the formulas

$$Q_1^{(p)} = \frac{\sum_{i=1}^{\lceil p \cdot J \rceil} L^{Inc}(i)}{\sum_{i=1}^{J} L^{Inc}(i)}, \quad Q_2^{(p)} = \frac{\sum_{i=1}^{\lceil p \cdot J \rceil} DP^{L\text{-sort}}(i)}{\sum_{i=1}^{J} DP^{L\text{-sort}}(i)}, \tag{6.2.2}$$

where p is the percentage corresponding to the portion of issuers which will cause a lower loss in the portfolio with respect to the other $(1-p)\,J$ issuers. We remark the fact that the index $Q_1^{(p)}$ represents the ratio of minimum possible loss in the portfolio $\sum_{i=1}^{J} L_j$ that would occur in the case that the portion p of issuers which would cause the lower loss in the portfolio will face the default event. Respectively, $Q_2^{(p)}$ is the concentration ratio of the default probabilities of the issuers which would determine lower losses in the portfolio.

In the regression we consider as regressors the concentration indexes $Q_1^{(90\%)}$ $Q_2^{(75\%)}$, i.e. through linear regression the model would be

$$y = \beta_0 + \beta_1\, Q_1^{(90\%)} + \beta_2\, Q_2^{(75\%)} + \epsilon\,, \tag{6.2.3}$$

where the slopes β_1 and β_2 and the intercept β_0 are the unknown parameters, the normalized quantile y is the dependent variable, the rate index, Q_1 and Q_2 are the independent variables and ϵ represents the estimation error. Namely, the underlying idea is that a general portfolio satisfying the characteristics arising from practical experience can adequately be represented by these two concentration indexes; and therefore, for high-dimension portfolios, it suffices to perform the regression for $Q_1^{(p_1)}$ and $Q_2^{(p_2)}$, for proper p_1 and p_2.

The choice of such a percentage p for $Q_1^{(p)}$ is embedded in the hypothesis of concentration above the construction of L. To select such p we simulate several curves $Q_1 : [0,1] \to [0,1]$, $p \mapsto Q_1^{(p)}$, as shown in Figure 6.4, and we consider the

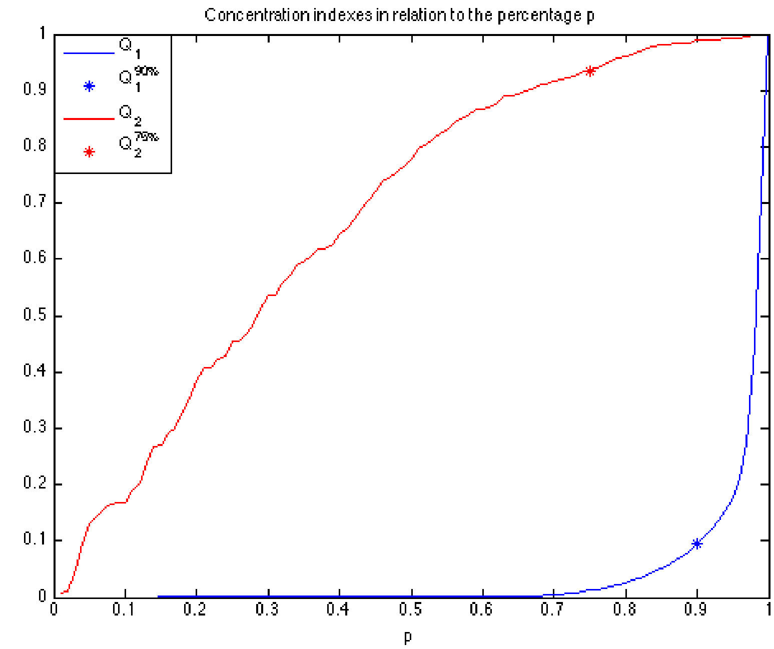

Figure 6.4: Plot of the concentration indexes of a simulated portfolio.

tangent lines $T_{\bar{p}}$ to the graph described by Q_1 in different points $\bar{p} \in (0,1)$

$$T_{\bar{p}}(x) = \frac{\mathrm{d}Q_1^{(p)}}{\mathrm{d}p}(x-\bar{p}) + \bar{p}\,. \tag{6.2.4}$$

In particular we consider the slopes of the tangent lines and it turns out that for $p \in (0, 0.85]$ these slopes are close to zero and far less than 1, which imply that the variations $\frac{\Delta Q_1^{(p)}}{\Delta p}$ are exiguous and it means that choosing $\bar{p} = 0.1$ or $\bar{p} = 0.7$ would result in $Q_1^{(\bar{p})}$ very close to each other; on the contrary for $\bar{p} \in [0.95, 1)$ the slopes are too large and therefore different simulations could give values of $Q_1^{(\bar{p})}$ too distant. For these reasons we choose a middle value $p = 0.9$, so that $Q_1^{(p)}$ is appropriate and meaningful for the regression. We follow the same theory to choose the percentage p for $Q_2^{(p)}$. This time the only care needed is p not being close to 1 since the issuers with higher L have very poor DP, so the variations $\frac{\Delta Q_2^{(p)}}{\Delta p}$ would be close to zero and the values of the regressors would be not very significant for the regression.

Figure 6.4 represents the concentration indexes for a portfolio simulated as explained previously, the chosen $Q_1^{90\%}$ and $Q_2^{75\%}$ are marked.

A linear regression model could be inappropriate to estimate the rate y. In fact, it would not consider the upper and lower boundaries, hence, to avoid predictions outside the interval $[0,1]$, we apply the *Logistic Quantile Regression*(LQR). There are several application of the LQR to systemic risk analysis, see [108] for a study concerning the forecast of recovery rates. Moreover, LQR is widely used in many statistical fields, as, e.g., in Economics, as well as in Ecology, Meteorology, Biomedical sciences, etc., see, e.g., [25, 75], and references therein.

Hence through the logistic formula for the dependent variable y

$$\mathrm{logit}(y) = \log\Big(\frac{y}{1-y}\Big)$$

we bound the estimated values between 0 and 1. Inverting the logit function we get the regression model as

$$y \sim \frac{1}{1 + \exp\big(-\big(\beta_0 + \beta_1 Q_1^{(90\%)} + \beta_2 Q_2^{(75\%)} + \epsilon\big)\big)} . \tag{6.2.5}$$

We start considering the simulation of $n = 50$ portfolios composed by shares of $J = 12$ different issuers. At this first step n does not play a specific role, we just use $n = 50$ as it is enough for an estimation of the 2 parameters in the regression. Each of them is simulated by the procedure explained in the previous subsection. So we compute n couples of concentration indexes $Q_1^{(90\%)}$ and $Q_2,^{(75\%)}$, and rate indexes y, and through the logit regression model we estimate the parameters β_0, β_1 and β_2.

Once the parameters β_0, β_1 and β_2 are estimated, they can be used to estimate the value of the rate y^* corresponding to the 99.9% worst case for the portfolio P^* with dimension J^* as

$$y^* = \mathrm{logit}^{-1}\big(-\widehat{\beta}_0 - \widehat{\beta}_1 Q_1^{(90\%)^*} - \widehat{\beta}_2 Q_2^{(75\%)^*}\big) , \tag{6.2.6}$$

where $Q_1^{(90\%)^*}$ and $Q_2^{(75\%)^*}$ are the concentration indexes for P^* and $\widehat{\beta}_0$, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ denote the estimates for the regressor coefficients. Consequently the estimated Loss for P^* is

$$\mathrm{IDR}^*(99.9\%) = y^* \cdot \sum_{i=1}^{J^*} L^*(i) , \tag{6.2.7}$$

where $L^*(i)$ is the loss determined by the default of the i^{th} issuer in the portfolio P^*.

6.2.1 Monte Carlo simulation and IDR estimation

Let us consider a portfolio P with J independent issuers. As before, a loss value $L(j)$ and a default probability $DP(j)$ are assigned to each issuer $j = 1, \dots, J$. The Monte Carlo (MC) method consists in the simulation of N default scenarios which can be summarized by the following steps:

1. Start simulating a uniform-randomly based matrix U, with values in $[0,1]^{J\times N}$, namely each of its elements $U(j,n)$, for $j \in \{1,\dots,J\}$ and $n \in \{1,\dots,N\}$, corresponds to the j-th issuer in the n-th scenario;

2. Then, for each scenario $n = 1\dots,N$, we have the j-th issuer default as soon as $U(j,n) < DP(j)$;

3. The loss in the portfolio P, for the n-th simulated scenario, is given by the sum of the $L(j)$ such that $U(j,n) < DP(j)$, i.e.

$$Loss^{MC}(n) = \sum_{j=1}^{J} L(j)\, \mathbb{1}_{\{U(j,n)<DP(j)\}}, \quad \text{for } n = 1,\dots,N. \tag{6.2.8}$$

We use the empirical quantile as the MC estimate of the theoretic quantile, i.e. the estimated $\mathrm{IDR}^{MC}(99.9\%)$ is the $(N/1000)$-th greatest $Loss^{MC}$.

By construction, we have that the random variable $Loss$ is such that

$$\mathbb{P}\big(Loss \le \mathrm{IDR}(99.9\%)\big) \approx 99.9\%\,, \tag{6.2.9}$$

where IDR(99.9%) is the 99.9% percentile that we want to estimate. Therefore the sample $p = 99.9\%$ quantile $\mathrm{IDR}^{MC}(p)$ satisfies the following equation

$$\hat{\sigma}_N^2 = Var(\mathrm{IDR}^{MC}) \approx \frac{p\,(1-p)}{(N+2)\, f^2(\mathrm{IDR})} + O(1/N^2)\,, \tag{6.2.10}$$

with f probability density function (PDF) of $Loss$, see, e.g., [34]. In Figure 6.5 is plotted f, the PDF estimated from the simulated data $Loss_1^{MC},\dots,Loss_N^{MC}$, with N =1,000,000. The estimation is based on a normal kernel function. We estimate the confidence interval with 99.9% confidence level as

$$CI_N^{MC}(99.9\%) = \mathrm{IDR}^{MC} \pm 3.0902\,\hat{\sigma}_N. \tag{6.2.11}$$

6.2.2 Comparison between heuristic model and exact algorithm

We consider $J^* = 18$ as the dimension of the portfolio target. Our aim is to compare the HR model with the exact algorithm and the MC model and see how far is the

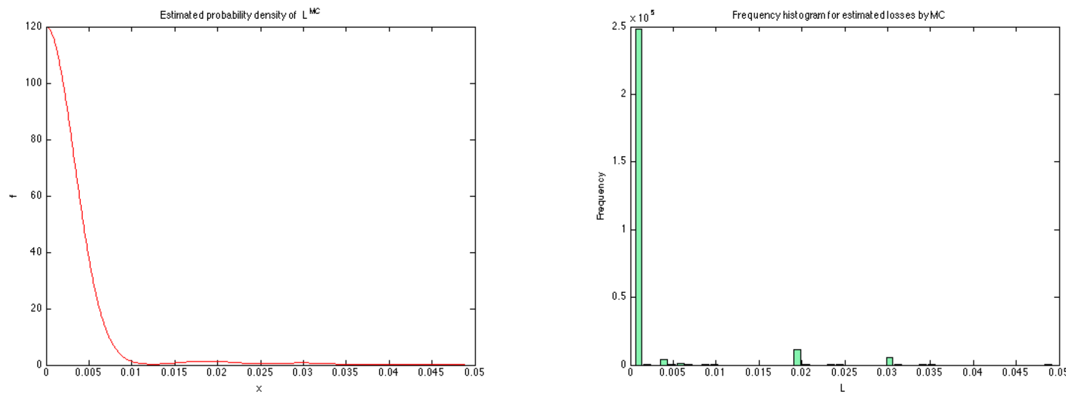

Figure 6.5: Estimated PDF of *Loss* from the MC simulated data represented in the adjacent frequency histogram.

prediction for the IDR obtained by the HR from the correct IDR obtained by the exact algorithm. So we simulate $n = 32$ regression portfolios with dimension $J = 12$ (at the end of this paragraph our choice for n and J will be explained).

The portfolios share the characteristics of beta distributed losses and default probabilities are inversely related to the correspondent losses. The choices of J^*, n and J are arbitrary, with the premise that $J < J^*$; but, moreover we have to explicit that n and J condition the time spent to realize the regression, since the regression model has complexity $T(J) = n \times O(2^J \ln(2^J)) \approx n \times O(2^J)$. From the logistic regression we get the parameters

$$\widehat{\beta}_0 = 3.2945\,,$$
$$\widehat{\beta}_1 = -5.6408\,,$$
$$\widehat{\beta}_2 = -0.8441\,.$$

Afterwards, these parameters are used to estimate the IDR of $N = 10$ portfolios target in order to compare them with the IDR computed by using the exact algorithm.

Alongside this we simulate $N^* = 10,000$ and $N^{**} = 100,000$ default scenarios for the MC method and we estimate the IDR and the CI, as shown in the previous paragraph.

In Figure 6.6 are plotted the rate values y^* obtained by the exact algorithm and the HR model, both corresponding to the N Portfolios Target, and in Table 6.1 are shown the obtained values compared with the CI resulting from the MC simulations. The order of magnitude concerning execution time and accuracy of the iterations of the HR and the MC model are not clearly comparable, since HR is much faster,

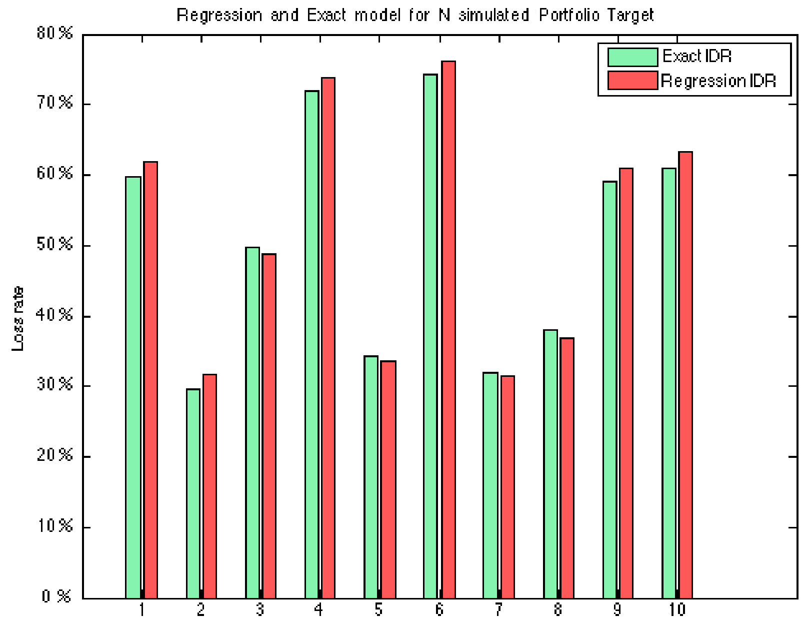

Figure 6.6: Histogram of the comparison between the values obtained with the exact algorithm and the HR model.

while MC is usually more accurate, especially for a great number of MC simulations N^{MC}. As to provide a meaningful comparison, we have simulated 20 portfolios with dimension $J^* = 18$. Figure 6.7 represents the execution times in a logarithmic scale and the estimation errors that we got from the HR method and the MC simulations for $N^{MC} = 10^3, 10^4, 10^5$; despite MC for $N^{MC} = 1000$ is not very meaningful, we considered it in order to compare it with HR and have a similar order of magnitude for the error. We point out the fact that even though the real distribution of the IDR estimated by MC is not known, we build the order statistics in order to estimate the probability density relying on a normal kernel function. Then the estimate of the PDF was exploited to estimate the variance of IDR^{MC}, and enable us to construct the CI and compare them with the errors of the HR method.

In order to compare the time consumption, we remark the fact that to compute the estimation for the regression coefficients β_0, β_1 and β_2 it took 28 seconds, and less than a millisecond for the prediction of the IDR for each portfolio target, i.e. by applying the HR model. Instead, the evaluation of the IDR took approximately a minute. The computation time for the simulation with MC depends on the dimension J^* and on the number of simulations N^{MC}; the times listed in Table 6.2 have to be compared with the millesimal time required by the HR model.

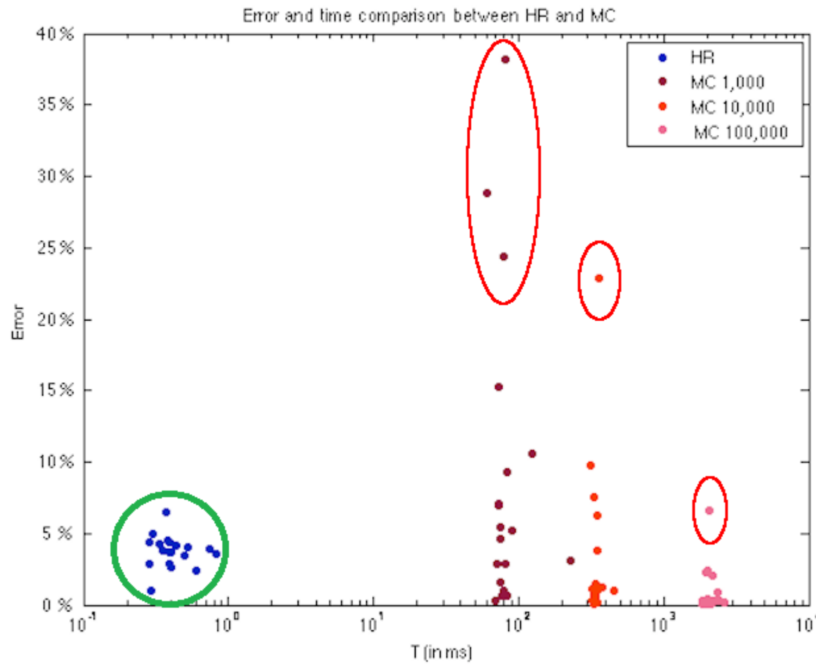

Figure 6.7: Comparison between HR and MC model that relates the execution times T and the estimation errors E. The HR error is given by the absolute value of the difference between the exact IDR and the estimated IDR, while the MC error is equal to the width of the estimated confidence interval, i.e. equal to 6.1804 $\hat{\sigma}_N$. We marked with a red circle the outliers of MC, and stress the fact that even for a MC with $N^{MC} = 10^4$, which is almost equal to $2^{J^*} = 262144$, there are chances of outliers.

Ex. IDR	HR IDR	Delta	$\hat{\sigma}^{MC}_{N^*}$	$CI^{MC}_{N^*}(99.9\%)$	$\hat{\sigma}^{MC}_{N^{**}}$	$CI^{MC}_{N^{**}}(99.9\%)$
59.83%	61.89%	-2.06%	0.24%	[59.21%, 60.69%]	0.05%	[59.73%, 60.03%]
29.60%	31.58%	-1.97%	0.87%	[28.90%, 34.28%]	0.20%	[29.53%, 30.76%]
49.76%	48.77%	0.99%	0.23%	[49.20%, 50.63%]	0.06%	[49.51%, 49.89%]
71.85%	73.77%	-1.92%	0.15%	[71.59%, 72.52%]	0.03%	[71.76%, 71.95%]
34.23%	33.52%	0.71%	0.09%	[33.71%, 34.26%]	0.01%	[34.21%, 34.28%]
74.29%	76.09%	-1.79%	0.50%	[73.45%, 76.54%]	0.12%	[74.02%, 74.76%]
31.93%	31.47%	0.46%	0.01%	[31.90%, 31.97%]	0.004%	[31.93%, 31.95%]
37.99%	36.87%	1.11%	0.54%	[36.12%, 39.46%]	0.18%	[37.24%, 38.35%]
59.04%	60.97%	-1.92%	0.28%	[58.62%, 60.11%]	0.07%	[58.87%, 59.30%]
60.82%	63.36%	-2.53%	0.02%	[60.78%, 60.91%]	0.003%	[60.82%, 60.83%]

Table 6.1: IDR resulting from the HR model and the exact algorithm for a $J^* = 18$ dimensional portfolio, where Delta = HR IDR − Exact IDR. The $\hat{\sigma}_{N^*}$ and $\hat{\sigma}_{N^{**}}$ are estimated by equation (6.2.10) with $N^* = 10,000$ and $N^{**} = 100,000$, the CI are centered in the IDR values obtained by the MC estimation.

J^*	MC time	HR time	$\text{MC}_{\text{time}}/\text{HR}_{\text{time}}$
15	3.83s	0.002s	1900
20	4.96s	0.004s	1250
25	5.86s	0.004s	1450
30	6.43s	0.003s	2150
40	8.64s	0.007s	1250

Table 6.2: Times spent to evaluate the IDR with the MC and HR methods and for different portfolio dimensions J^* (the parameters $\beta_0, \beta_1, \beta_2$ were estimated once and for all). In the last column we indicate how many times the HR method results faster than the MC method.

Moreover, increasing the dimension of the problem, the time to estimate the IDR with the HR approach remains almost the same, indeed previously we saw that the computational complexity is $T(J^*) \approx n \times O(2^J)$, i.e. it depends on the chosen regression parameters n and J, instead of J^*. On the other hand, the time to evaluate the IDR with the exact approach of Section 6.1 increases exponentially. This is due to the fact that the computational complexity for the exact algorithm is related to the input parameter J^*, since the mean complexity of the best possible sort of a k-dimensional vector is $O(k\ \ln(k))$, see [74, sec. 5.4, pp. 248379]. Therefore, the resulting complexity is

$$T(J^*) = O(2^{J^*}\ \ln(2^{J^*})) = O(2^{J^*}\ J^*\ \ln(2)) \approx O(2^{J^*})\,. \tag{6.2.12}$$

We have this order of computational complexity since we are using an exact algorithm that sorts the vector of the 2^{J^*} possible loss outcomes, then it computes the CDP until it finally stops when it reaches the desired percentile.

An alternative approach for the exact algorithm is represented by the scanning of the tree of cases, i.e. an algorithm that firstly computes the greatest loss L_1, and its CDP, then the second greatest loss L_2 and its CDP, the third and so on, until the percentile 0.1% is reached. But for the choice of structuring the portfolios as explained at the beginning of this chapter, it is necessary to consider approximately the half of all possible combinations, and therefore it takes even more time than the previous exact algorithm.

Let us underline that we have decided to consider $J = 12$ and $n = 32$, with respect to $J^* = 18$, mainly because of the computational effort required by the HR and exact method. In particular, to make the HR more efficient with respect to the

J	n: 32	n: 64
12	29”	53’
13	54”	1’ 56”
14	1’ 48”	3’ 37”
15	3’ 32”	7’ 17”
16	7’ 12”	13’ 50”

Table 6.3: The time to compute the regression does not depend on the Portfolio target P^*, and in particular not on J^*. This table reports the times spent to compute the regression for different J and n. We remark the fact that the regression time is a set-up which can be estimated once a month, and therefore since the meaningful time for the HR method is the one reported in the third column of table 6.2, for high Portfolio dimensions, the time saved is considerable.

exact calculation, i.e. with smaller computational complexity

$$O(\text{HR}) < O(\text{exact}) ,$$

we need $n \times O(2^J)) < O(2^{J^*})$, hence

$$J^* - \log_2(n) > J ,$$

see equation (6.2.12) for the complexity of the exact algorithm. Therefore, since $n = 32$ is a sufficient sample size, for $J^* = 18$, J has to be at least less than 14. In Table 6.3 are shown the times spent to compute the regression model for various J and n, these times have to be compared with the time spent to execute the exact algorithm.

6.2.3 The computational workflow - A summary

In the application perspective, the accuracy and the execution time are very important, but we also need to have a clear description of the algorithm, its set-up complexity, the maintenance effort and so on.

Generally the “long” bank bond portfolios change quite slowly, hence the concentration and correlation measures between the exposures are very smoothly over time. Hence, we can split the workflow in 3 different levels. A first level is the general (near static) set-up, i.e. the definition of the general parameters, such as the number of simulation N, the regression technique, the size J for the benchmark portfolio, the α levels for Q_1 and Q_2, etc.

The second level, with a periodic update, is given by the core parameters estimation, i.e. the IDR calculation for benchmark portfolios and the parameters (β_1, β_2) estimation by the regression procedures.

Finally the third level is the execution task, in other words the calculation of the estimate IDR, given the current actual portfolio parameters and the estimated β's.

As concerns the frequency of the above processes, the general set-up is usually triggered by any top management (risk control, risk committee) new guidelines or yearly. The periodic update could be run monthly or quarterly according to the portfolio dynamics. Finally the execution process is related to the bank regulatory constraints. If the bank has validated the internal model for the specific market risk, e.g. following a specific solution as the one underlined in [23] a daily calculation is usually required. Otherwise the bank can set its own reporting frequency.

To summarize, the proposal can fit quite easily the internal usual bank procedures and processes.

6.3 Extension to the correlated case and possible approaches

Until now only the case with uncorrelated issuers was considered. In the previous sections of this chapter the outside-diagonal elements of the correlation matrix were set equal zero. Although this setting is useful to be the starting point to deal with the heuristic approaches to the IDR estimation problem, from a financial point of view this is unrealistic.

In the non-independent portfolio case, i.e. removing the issuers' independence assumption, also the elements outside the diagonal are non-zero. Correlations are not observable from the market, and depend on the underlying assumptions of the assets values, such as the Black and Scholes model, from which they can be obtained. However from a theoretical point of view the correlations are estimated from the CDS spreads of the the issuers: but usually this is impracticable also for the shortage of data. So often they are estimated from the equity prices. Furthermore computing the IDR for dependent portfolios is more complex, and in particular unfeasible even for portfolios with not very huge dimensions due to the great execution time.

Therefore our aim is to build an adequate structure to apply an HR model. The technical problem underlying the new HR model is that we would not have to estimate only the coefficients β_0, β_1, β_2 of the previous (uncorrelated) model, but also all the coefficients concerning the correlations between the issuers. Since this is impracticable, we have to find less variables that properly represent the correlations between the issuers.

Before explaining our proposal, let us summarize what the FRTB regulations claims for the IDR model parameters calibration. The general point is that the high confidence level for the measure, i.e. 99.9%, along with the rather long time horizon, i.e. 1 year, does not allow for an easy IDR-back test procedure for both the default events of the financial instruments issuers and for the related correlations. Practically, this is not feasible, hence in the FRTB new framework some flexibility is allowed, see [13]. In particular, we have the following main points to be considered:

- Correlations should be estimated over a 10 years periods, covering a stress period, hence allowing to consider increases of correlations during the financial crisis
- Correlations must be inferred from spreads or equity prices, with some proxies/benchmark where not available
- Banks must have clear policies and procedures that describe the correlation calibration process
- Previous correlations must be based on objective data
- A bank must validate its modeling approach for such correlations, namely it has to show that the provided *internal* method is appropriate for the detained portfolio, also with respect to the chosen systematic risk factors and associated weights.

We would also like to point out the remark written by BCBS about the DRC validation, namely "*Accordingly, the validation of a DRC model has to rely more heavily on indirect methods including, but not limited to, the stress tests, sensitivity analysis and scenario analysis, hence to assess both its qualitative and quantitative efficiency and robustness, particularly with regard to the models treatment of concentrations*".

To be more explicit, banks must have sound procedures and processes to assess, build and update their models with proper documentation, even if, due to the very challenging goal of the IDR estimation, some flexibility is permitted.

We note that usually portfolios are comprised of issuers that are correlated to each other in a similar way (this feature is due to the geographic consistence of portfolios). Therefore the outside-diagonal correlation matrix elements don't vary very much and we can consider the mean value of them as one of the regressors of the new HR model, let us call it $\widetilde{\sigma}$. Moreover in the regression we add the regressor D, corresponding to the index of dispersion of the outside diagonal elements. So, for each portfolio, simultaneously to the computation of $Q_1^{(90\%)}$ and $Q_2^{(75\%)}$, obtained as concentration indexes of L and DP respectively, we need to compute $\widetilde{\sigma}$ and D.

Therefore the new linear regression will be

$$y = \beta_0 + \beta_1 \, Q_1^{(90\%)} + \beta_2 \, Q_2^{(75\%)} + \beta_3 \, \widetilde{\sigma} + \beta_4 \, D + \epsilon \,, \tag{6.3.1}$$

where the dependent variable y is the normalized quantile, the intercept β_0 and the slopes β_i for $i = 1, 2, 3, 4$ are the unknown parameters, the rate indexes Q_1 and Q_2, the mean correlation $\widetilde{\sigma}$ and the dispersion index D are the independent variables, and ϵ represents the estimation error.

Part IV

Portfolio optimization for Hawkes-Merton models with transaction costs

I will tell you the secret to getting rich on Wall Street. You try to be greedy when others are fearful. And you try to be fearful when others are greedy.

Warren Buffett (1930-)

Introduction

The optimal consumption-investment problem in continuous time, pioneered by Merton [89] has been extensively investigated in many different model settings. In the original approach, the risky asset dynamics was described by a geometric Brownian motion. This assumption has been proved to be quite naive during the last 30 years. The model proposed by Merton for the portfolio optimization problem, and later adopted by Merton himself and by Black and Scholes in their historical paper on European option pricing [20], has been refined and modified over the years in order to take into account more realistic features of asset prices. Extensions have been proposed in several directions: some models include a stochastic dynamics for volatility, some include jumps into the asset price dynamics in order to describe sudden and unexpected price variations, difficult to explain by Gaussian fluctuations, some model include transaction costs. The portfolio optimization problem for stochastic volatility models was investigated by Kraft [77] in a Heston-type setting; Zeng and Taksar [118] studied the case of power utility by providing an explicit solution for the optimal portfolio problem still in the Heston setting.

As far as models with jumps are concerned, among the first contributions provided to the optimal consumption-investment problem we mention the papers by Jeanblanc and Pontier [81], by Aase [1] and by Aït-Sahalia et al. [4].

The portfolio optimization problem with proportional transaction costs has been extensively investigated in the paper by Davis and Norman [44], where an exhaustive qualitative analysis is performed of the optimal investment strategy. Shreve and Soner [109] extended this result in order to include short positions and they characterized mathematically the value function by using the viscosity approach. Øksendal and Sulem [112] discussed the portfolio optimization problem in a market model where both fixed and proportional transaction costs appear.

Up to some extent, all these papers show the robustness of results obtained by Merton, since the optimal policy obtained is similar to the policy computed by him, suitably modified in order to take into account the different particular features with a direct effect. For instance, the presence of jumps increases the variance of the stock and reduces then the exposition of the optimal portfolio, while transaction costs reduce dramatically the number of portfolio reallocations.

When transaction costs are included in modeling financial markets, the optimal portfolio problem is not only relevant *per se*, but also in view of derivatives pricing. By applying the so-called Utility Indifference Pricing approach, introduced by Hodges and Neuberger [80], Davis et al. [45] provided a suitable valuation procedure of contingent claims in models with transaction costs based on portfolio optimization strategy.

In this part of the thesis, a financial market is considered, composed by two investment opportunities: a *risk-free asset*, also called money market or Government bond or simply bond, which evolves at a risk-free interest rate, and a *risky asset*, also referred to as stock or share, subject to a significant degree of volatility and a jump component driven by a Hawkes process.

We now detail the contribution these chapters by comparing the usual constant intensity case (Poisson) to our self-exciting structure (Hawkes). The variance of the stock is explained by both the Brownian and the jump part. In particular, the stock variance turns out to be an increasing function of the jumps intensity; and, since the investors are risk-adverse, the slope of Merton line, see e.g. [89], and of the two buy-sell lines delimiting the no-transaction region, see e.g. [44] , [109], are decreasing with the variance. In other words these three lines spin downwards for an increased jump intensity. We shall call the Merton and no-transaction lines (UNT, Upper No Transaction and LNT, lower No Transaction lines respectively) the three decision lines.

In a Poisson framework, a jump does not affect the intensity of future jumps, and therefore, after a jump in the stock dynamics take place, the position of the three decision lines stays unchanged. As a consequence after a negative (resp. positive) jump the optimal strategy is to do nothing or to buy (resp. sell) the asset depending whether the jumps is smaller or larger than a threshold given by the buy (resp. sell) line. Therefore, the result of the pioneering model on transaction costs by Davis and Norman [44] is consistent with the Poissonian jumps framework, see e.g. [113], since in both framework the no-transaction region is static.

Instead, in a Hawkes framework, when a jump occurs, the intensity will increase dramatically. The main consequence is that the three decision lines turn clockwise brutally due to the increase of future variance. Conversely, in a period of lack of jumps, the same lines turn anticlockwise slowly due to the smooth decreasing of intensity between two jumps.

As a consequence, the symmetry of the original model is broken and in particular the jump sign becomes significant. We first focus on the effect of a positive jump: since after a jump the lines turn clockwise, the policy for a Hawkes framework is mainly unchanged with respect to the Poisson framework, and probably magnified, i.e. the investor has to sell the risky asset in the case of a jump overflowing the no-transaction region. Conversely, this policy is no longer optimal when negative jumps are considered, since a third zone can appear as detailed in the Figure 6.8. That is, if the jump size is smaller than a threshold the decrease of the ratio between risky and risk-free assets could be more than compensated by the increase on the portfolio variance. In other words, the direct effect of a negative jump could be smaller than the feedback effect related to self exciting property of Hawkes process.

Globally the effect is that it is possible that rational investors have to sell risky asset after a negative jump. In contrast, positive jumps do not have this snowball effect. This two-sided phenomenon is mainly motivated by the risk aversion of the investors. Indeed, in a risk-lover case, a similar analysis will give birth to a snowball effect only for positive jumps.

Moreover we highlight that the self exciting property of our model is self-produced only when negative jumps are concerned. Positive jumps could exist but are pure (compound) Poisson.

We now focus on a period between two jumps. As time goes up the intensity of the Hawkes process decreases toward its basic value. The result is that the variance of the risky asset decreases smoothly with the time. As highlighted before the effect is that the lines turn anticlockwise but this movement is regular. The effect is that investors along the buy line will progressively increase their exposure on the risky asset, not due to a change on the asset's drift, but only on the reduction of its variance risk, which means that between two jumps the asset price increases supported by investors buying for endogenous variance reduction reasons and not for an exogenous change of drift.

This fourth part of the thesis is organized as follows: in Chapter 7 we are going to present our modeling framework. In Section 7.2 we prove some properties of the value function resulting from this problem and in Section 7.3 we provide a dynamic programming principle. In Section 7.5 we investigate existence, uniqueness and regularity of the solution for the problem formulated in the previous sections. In Section 8.1 we show that, under suitable hypothesis, the dimension of previous problem can be reduced by a similarity argument, and we provide a qualitative description of the reduced problem solution in Section 8.3. In Section 8.2 we present a numerical scheme in order to obtain an approximate solution of the previous problem.

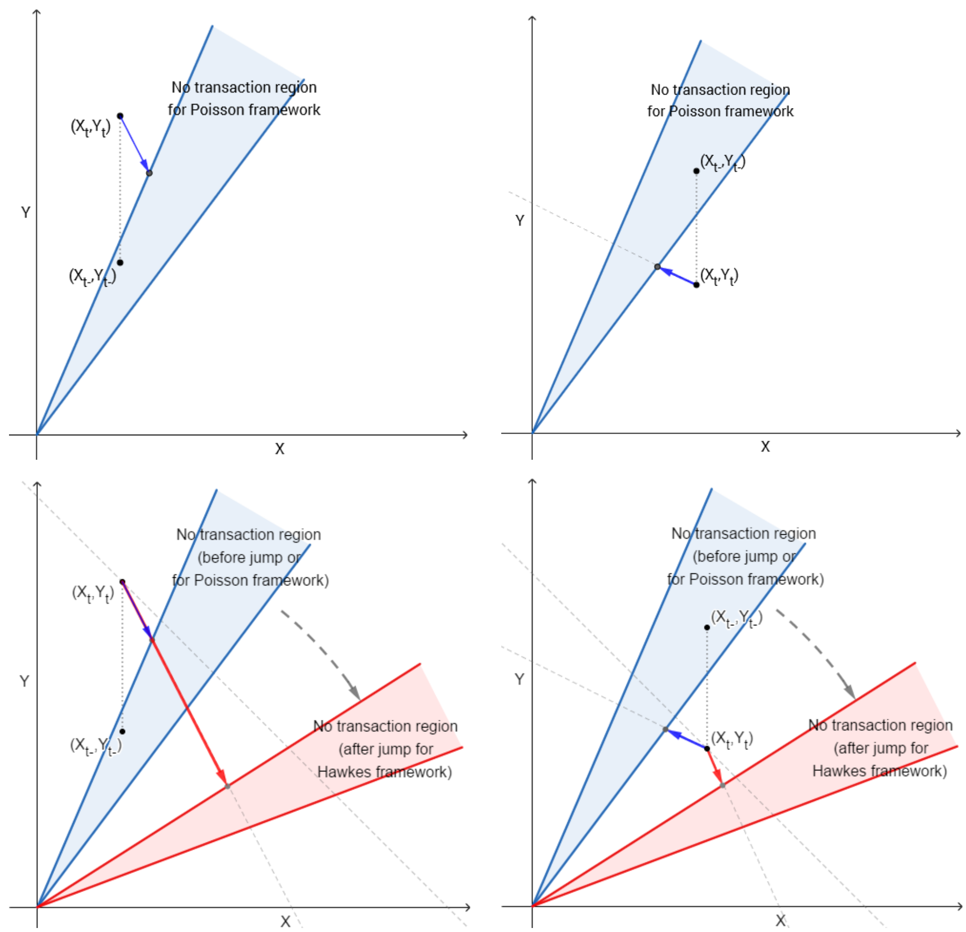

Figure 6.8: Top figures represent the Poisson framework, while bottom figures the Hawkes framework. In the left figures positive jumps are considered, while in the right figures negative jumps. The top figures are consistent with the efficient-market hypothesis, that is “fearful when others are greedy and greedy only when others are fearful” quoting W. Buffet. On the other hand, the bottom figures represent the influence of the emotional states influencing market behaviors, and which may cause unpredictability and volatility in the stock market.

Chapter 7

Merton-Hawkes model with proportional transaction costs

This chapter deals with the mathematical characterization of the related optimal policy. In the same spirit of Davis and Norman [44] and Shreve and Soner [109], we will show that the Merton line is replaced by two lines identifying three regions, one where is optimal to buy, another to sell and a third one where is optimal to wait due to the costs.

7.1 The model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered, complete probability space, with right-continuous filtration, supporting a Brownian motion and a self-exciting point process, and consider a market consisting of two investment opportunities: a risky asset $\{S_t\}_{t\in\mathbb{R}^+}$ and a risk-free asset $\{S_t^0\}_{t\in\mathbb{R}^+}$, evolving according to the following equations

$$\begin{cases} \mathrm{d}S_t &= S_t \left(\alpha \, \mathrm{d}t + \sigma \, \mathrm{d}W_t + \int_{-1}^{+\infty} \eta \, \mathcal{N}(\mathrm{d}t, \mathrm{d}\eta) \right) \\ \mathrm{d}S_t^0 &= r \, S_t^0 \, \mathrm{d}t \end{cases} \tag{7.1.1}$$

where $\mathcal{N}$ is the Hawkes process, W is a Brownian motion $\mathcal{F}_t$-adapted, $\alpha, r \in \mathbb{R}^+$ are parameters representing the mean rate of return of the risky asset and the risk-free rate respectively, such that $\alpha > r$, $\sigma \in \mathbb{R}^+$ is the volatility, while $\mathrm{d}N_t = \int_{-1}^{+\infty} \mathcal{N}(\mathrm{d}t, \mathrm{d}\eta)$ is a Hawkes process $\mathcal{F}_t$-adapted with Lévy measure ν with negative expectation and with exponentially decaying intensity λ_t, i.e. the intensity process $\{\lambda_t\}$ satisfies:

$$\mathrm{d}\lambda_t = a\,(b - \lambda_t)\,\mathrm{d}t + \kappa\,\mathrm{d}N_t\,,$$

or equivalently, by Itô lemma

$$\lambda_t = b + e^{-a\,t}\,(\lambda_0 - b) + \kappa \int_0^t e^{-a\,(t-s)}\,\mathrm{d}N_s\,,$$

where a is the speed of mean reversion of the intensity to its long-run mean b, κ is the increase in the intensity generated by a jump, and λ_0 is its starting value, with $\lambda_0 \geq b$. We remark the fact that the class of processes to which N_t belongs are non-Markovian extensions of Poisson processes, but at the same time (N_t, λ_t) is a continuous-time Markov process.

Moreover, consider an investor holding a portfolio, starting with a positive position x invested in the risk-free asset and a positive position y invested in the risky asset, and assume that she is able to consume the amount invested in the risk-free asset and to transfer its capitals from an investment to another paying proportional transaction costs. Therefore let us introduce the instantaneous consumption process C, assumed to be non-negative and integrable on each finite time interval, the processes L representing the cumulative amount of risk-free asset sold in order to buy risky asset, and M the process representing the cumulative amount of risky asset sold in order to buy risk-free asset. Both L and M are assumed to be non-negative, non-decreasing and càdlàg.

The portfolio value can be represented continuously in time by the couple (X_t, Y_t), for $t \geq 0$ and starting at $X_0 = x$, $Y_0 = y$, with (X, Y) representing the amount of capital invested in the risk-free asset and in the risky asset, respectively, and evolving according to the following stochastic differential equations

$$\mathrm{d}X_t = (r\,X_t - C_t)\,\mathrm{d}t + (1 - \mu_{M})\,\mathrm{d}M_t - \mathrm{d}L_t\,, \tag{7.1.2}$$

$$\mathrm{d}Y_t = \alpha\,Y_t\,\mathrm{d}t + \sigma\,Y_t\,\mathrm{d}W_t + Y_{t^-} \int_{-1}^{+\infty} \eta\,\mathcal{N}(\mathrm{d}t, \mathrm{d}\eta) + (1 - \mu_{L})\,\mathrm{d}L_t - \mathrm{d}M_t\,, \tag{7.1.3}$$

where $\mu_{M} \in [0, 1)$ is the proportional transaction cost associated to transfer a unit of risky asset to $(1 - \mu_{M})$ units of risk-free asset, while $\mu_{L} \in [0, 1)$ corresponds to the relocation of the investment in the risk-free asset to the risky asset.

Because of the positive and negative jumps, we define the solvency region as in [44], i.e. $\mathcal{S} := \{(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+\}$, and denote by $\mathcal{A}(x, y)$ the set of admissible consumption/investment policies for $(x,y) \in \overline{\mathcal{S}}$, defined as the set of (C,L,M) such that $(X_t, Y1_t)$, given by (7.1.2) and (7.1.3), stays in $\overline{\mathcal{S}}$ for all $t \geq 0$. In [109] the investments dynamics were assumed to be continuous, and therefore also short positions on both the risk-free and the risky asset were allowed, as long as the agent was able to cover his short positions of one investment with the long position of the other. In our case we need to restrict the solvency region to the 1st quadrant, because at any time a jump can occur and we are considering risk-adverse investors.

For $p \in (0,1)$ fixed, let us introduce the agent's utility function U_p defined as $U_p = c^p/p$, for all $c \geq 0$ and a positive discount factor β, and we assume that the agent aims to maximize the expected value of the discounted value of its utility function. Therefore the maximization problem has the following value function

$$v(x,y,\lambda) = \sup_{(C,L,M)\in\mathcal{A}(x,y)} \mathbb{E}\left[\int_0^\infty e^{-\beta\, t}\, U_p(C_t)\,\mathrm{d}t\right], \quad \forall (x,y,\lambda) \in \overline{\mathcal{S}} \times [b,+\infty). \quad (7.1.4)$$

Notice that the utility function is concave, and therefore the agent is assumed to be risk adverse. This point will be better remarked later when the change in intensity will be considered.

7.2 Basic properties of the value function

Proposition 7.2.1 *The value function v defined by (7.1.4) is*

1. *concave on $\overline{\mathcal{S}}$;*
2. *p-homothetic with respect to the first two components, i.e. $v(\gamma\, x, \gamma\, y, \lambda) = \gamma^p\, v(x,y,\lambda)$, $\forall (x,y,\lambda) \in \overline{S} \times [b,+\infty)$ and $\gamma > 0$;*
3. *strictly increasing with respect to the first two components.*

Proof.

1. By the linearity of (7.1.2) and (7.1.3) we have that linear combinations of two admissible policies for two different starting values of the couple (x,y) are still admissible policies for the linear combination of the starting values: i.e. for (x_1,y_1), $(x_2,y_2) \in \overline{S}$, $\gamma \in (0,1)$, $(C_1,L_1,M_1) \in \mathcal{A}(x_1,y_1)$, and $(C_2,L_2,M_2) \in \mathcal{A}(x_2,y_2)$ we have

$$\begin{aligned}&\big(\gamma C_1 + (1-\gamma)C_2, \gamma L_1 + (1-\gamma)L_2, \gamma M_1 + (1-\gamma)M_2\big)\\ &\qquad\qquad \in \mathcal{A}(\gamma x_1 + (1-\gamma)x_2, \gamma y_1 + (1-\gamma)y_2).\end{aligned}$$

Therefore

$$\begin{aligned} &v(\gamma\, x_1 + (1-\gamma)\, x_2, \gamma\, y_1 + (1-\gamma)\, y_2, \lambda) \\ &\qquad \geq \mathbb{E}\left[\int_0^\infty e^{-\beta\, t} U_p(\gamma\, C_1(t) + (1-\gamma)\, C_2(t))\,\mathrm{d}t\right] \\ &\qquad \geq \gamma\,\mathbb{E}\left[\int_0^\infty e^{-\beta\, t} U_p(C_1(t))\,\mathrm{d}t\right] + (1-\gamma)\,\mathbb{E}\left[\int_0^\infty e^{-\beta\, t} U_p(C_2(t))\,\mathrm{d}t\right], \end{aligned}$$

where the last inequality follows by the concavity of U_p. Then maximizing over the first term of the right hand side over $(C_1, L_1, M_1) \in \mathcal{A}(x_1, y_1)$ and the second over$(C_2, L_2, M_2) \in \mathcal{A}(x_2, y_2)$, we obtain property 1.

2. It is a direct consequence of the fact that $(C, L, M) \in \mathcal{A}(x, y)$ if and only if $(\gamma\, C, \gamma\, L, \gamma\, M) \in \mathcal{A}(\gamma\, x, \gamma\, y)$.

3. Let us consider $\delta \in \mathbb{R}^+$, then by simply applying the sub-optimal admissible strategy consuming the quantity δ at the starting time we have

$$v(x+\delta, y, \lambda) \geq v(x, y, \lambda) + U_p(\delta) > v(x, y, \lambda),$$

i.e. v is increasing with respect to the first component. For the increasing property with respect to the second component, consider a strategy that at the starting time transfers the quantity δ from the capital invested in the risky asset to the capital invested in the risk-free asset, and consumes the quantity $(1-\mu_M)\,\delta$, i.e.

$$v(x, y+\delta, \lambda) \geq v(x + (1-\mu_M)\,\delta, y, \lambda) \geq v(x, y, \lambda) + U_p((1-\mu_M)\,\delta) > v(x, y, \lambda).$$

□

We have the following lower bound for the value function.

Proposition 7.2.2 *For $\beta > r\,p$ and defining $C_* := \frac{\beta - r\,p}{1-p}$, for all $(x, y, \lambda) \in \overline{\mathcal{S}} \times [b, +\infty)$ we have*

$$v(x, y, \lambda) \geq \frac{1}{p}\, C_*^{p-1}\,(x + (1-\mu_M)\,y)^p.$$

Proof. Consider the admissible strategy consisting in transferring at the original time the capital invested in the risky asset in the investment in the risk-free asset

$$v(x, y, \lambda) \geq v(x + (1-\mu_M)\,y, 0, \lambda) = \sup_{C \in \mathcal{A}(x+(1-\mu_M)\,y, 0)} \int_0^\infty e^{-\beta\, t} \frac{C_t^p}{p}\,\mathrm{d}t\,. \qquad (7.2.1)$$

In particular our original problem (7.1.4) is reduced to a purely deterministic optimal problem. Moreover let us consider the admissible consumption strategy consisting in consuming continuously in time the proportional quantity $C_t = \gamma\, X_t$. Therefore (7.1.2) reduces to $X_t = (x + (1-\mu_M)\, y)e^{(r-\gamma)t}$, we have also the dynamics of C, and in particular we have a lower bound to (7.2.1):

$$v(x + (1-\mu_M)\, y, 0, \lambda) \geq \frac{(x + (1-\mu_M)\, y)^p}{p} \max_{\gamma\in(0,1)} \frac{\gamma^p}{\beta + p\,(\gamma - r)},$$

where $\gamma^* = \operatorname{argmax}_{\gamma\in(0,1)} \frac{\gamma^p}{\beta+p\,(\gamma-r)} = \frac{\beta - r\,p}{1-p} =: C_*$. □

7.3 Dynamic programming principle and related HJB equation

Let us formulate the Dynamic Programming Principle (DPP) for our problem:

$$v(x,y,\lambda) = \sup_{(C,L,M)\in\mathcal{A}_{[0,\tau]}(x,y)} \mathbb{E}\left[\int_0^\tau e^{-\beta\, t}\, U_p(C_t)\, \mathrm{d}t + e^{-\beta\,\tau}\, v(X_\tau, Y_\tau, \lambda_\tau)\right], \qquad (7.3.1)$$

for every $(x,y,\lambda) \in \overline{\mathcal{S}} \times [b,\infty)$ and where τ is a stopping time, possibly depending on $(C,L,M) \in \mathcal{A}_{[0,\tau]}(x,y)$ in (7.3.1).

Define the following integro-variational inequalities of dimension three, also called Hamilton-Jacobi-Bellman (HJB) equation:

$$\min\Big\{\beta\, v - Av - \mathcal{J}v - \widetilde{U}_p(v_x);\; v_y - (1-\mu_M)\, v_x;\; v_x - (1-\mu_L)\, v_y\Big\} = 0\,, \qquad (7.3.2)$$

where we have defined the following operators

$$Av(x,y,\lambda) := \frac{1}{2}\,\sigma^2\, y^2\, v_{yy}(x,y,\lambda) + \alpha\, y\, v_y(x,y,\lambda) + r\, x\, v_x(x,y,\lambda) + a\,(b-\lambda)\, v_\lambda(x,y,\lambda)\,,$$

$$\mathcal{J}v(x,y,\lambda) := \lambda \int_{-1}^{\infty} \Big(v(x,(1+\eta)y, \lambda+\kappa) - v(x,y,\lambda)\Big)\, \nu(\mathrm{d}\eta)\,,$$

$$\widetilde{U}_p(\widetilde{c}) := \sup_{c\in\mathbb{R}^+}\{U_p(c) - c\,\widetilde{c}\} = \frac{1-p}{p}\,\widetilde{c}^{\,p/(p-1)}\,.$$

In section 7.5 it will be proven that the value function defined in (7.1.4) is super- and sub- solution in viscosity sense for the HJB equation (7.3.2), and therefore after proving the uniqueness result we are allowed to study directly the HJB instead of the value function given by equation (7.1.4).

7.4 Continuity and upper-bounds to the value function

With the next result we are going to prove that supersolutions to the HJB equation (7.3.2) dominates the value function defined (7.1.4).

Proposition 7.4.1 *Let* $(x, y, \lambda) \in \overline{\mathcal{S}} \times [b, +\infty)$*, if* $\varphi \in C^2(\overline{\mathcal{S}} \times [b, +\infty))$ *is such that* $\varphi(0,0) = 0$*, increasing in the first two components, and*

$$\min\Big\{\beta\,\varphi - A\varphi - \mathcal{J}\varphi - \widetilde{U}_p(\varphi_x);\ \varphi_y - (1-\mu_{\scriptscriptstyle M})\,\varphi_x;\ \varphi_x - (1-\mu_{\scriptscriptstyle L})\,\varphi_y\Big\} \geq 0\,, \quad (7.4.1)$$

then we have $\varphi \geq v$*.*

Proof. Consider t such that

$$\theta_k \ \leq\ t\ <\ \theta_{k+1} \wedge \tau_{m,k} =: \widetilde{t}_{m,k}\,,$$

where $\{\theta_k\}_{k\in\mathbb{N}}$ is the sequence of arrival times in the Hawkes process, and $\tau_{m,k} := \inf\{t \geq \theta_k\ :\ Y_t \geq m \text{ or } Y_t \leq 1/m\}$. In what follows we will use the simplified notation $\varphi(t) = \varphi(X_t, Y_t, \lambda_t)$. Apply Itô formula to the discounted supersolution, to obtain

$$\begin{aligned}
e^{-\beta\,\widetilde{t}_{m,k}}\,\varphi(\widetilde{t}_{m,k}) &= e^{-\beta\,\theta_k}\varphi(\theta_k) + \int_{\theta_k}^{\widetilde{t}_{m,k}} e^{-\beta\,t}\big(-\beta\varphi(t) + A\varphi(t) - C_t\,\varphi_x(t) + \mathcal{J}\varphi(t)\big)\,\mathrm{d}t \\
&\quad + \int_{\theta_k}^{\widetilde{t}_{m,k}} e^{-\beta t}\sigma Y_t \varphi_y(t)\mathrm{d}W_t + \int_{\theta_k}^{\widetilde{t}^-_{m,k}} e^{-\beta\,t}\,\big((1-\mu_{\scriptscriptstyle M})\,\varphi_y(t) - \varphi_x(t)\big)\mathrm{d}L_t \\
&\quad + \int_{\theta_k}^{\widetilde{t}^-_{m,k}} e^{-\beta\,t}\,\big((1-\mu_{\scriptscriptstyle L})\,\varphi_x(t) - \varphi_y(t)\big)\,\mathrm{d}M_t - \int_{\theta_k}^{\widetilde{t}_{m,k}} e^{-\beta\,t}\,\mathcal{J}\varphi(t)\,\mathrm{d}t\,.
\end{aligned}$$

Since for all $c > 0$

$$\begin{aligned}
A\varphi + \widetilde{U}_p(\varphi_x) &= \sup_{c^*>0}\Big\{\frac{1}{2}\,\sigma^2\,y^2\,\varphi_{yy}(x,y,\lambda) + \alpha\,y\,\varphi_y(x,y,\lambda) + (r\,x - c^*)\,\varphi_x(x,y,\lambda) \\
&\qquad + a\,(b-\lambda)\,\varphi_\lambda(x,y,\lambda) + U_p(c^*)\Big\} \\
&\geq \frac{1}{2}\,\sigma^2\,y^2\,\varphi_{yy}(x,y,\lambda) + \alpha\,y\,\varphi_y(x,y,\lambda) + (r\,x - c)\,\varphi_x(x,y,\lambda) \\
&\qquad + a\,(b-\lambda)\,\varphi_\lambda(x,y,\lambda) + U_p(c)\,,
\end{aligned}$$

then by (7.4.1) and taking the expectation we obtain

$$\mathbb{E}[e^{-\beta\theta_k}\varphi(\theta_k)] \geq \mathbb{E}[e^{-\beta\widetilde{t}_{m,k}}\varphi(\widetilde{t}_{m,k})] + \mathbb{E}\left[\int_{\theta_k}^{\widetilde{t}_{m,k}} e^{-\beta\, t}U_p(C_t)\mathrm{d}t\right] + \mathbb{E}\left[\int_{\theta_k}^{\widetilde{t}_{m,k}} e^{-\beta\, t}\mathcal{J}\varphi(t)\mathrm{d}t\right],$$

and by Fatou's lemma

$$\begin{aligned}\mathbb{E}[e^{-\beta\,\theta_k}\varphi(\theta_k)] \geq \mathbb{E}[e^{-\beta\,\theta_{k+1}}\,\varphi(X_{\theta_{k+1}}, Y_{\theta_{k+1}^-}, \lambda_{\theta_{k+1}^-})]\\ + \mathbb{E}\left[\int_{\theta_k}^{\theta_{k+1}} e^{-\beta\, t}U_p(C_t)\,\mathrm{d}t\right] + \mathbb{E}\left[\int_{\theta_k}^{\theta_{k+1}^-} e^{-\beta\, t}\,\mathcal{J}\varphi(t)\,\mathrm{d}t\right].\end{aligned}$$

Therefore since $\int_0^t \mathcal{J}\varphi(s)\,\mathrm{d}s$ is the compensator of the pure jump process

$$\sum_{k=0}^{N_t} e^{-\beta\,\theta_k}(\varphi(\theta_k^-) - \varphi(\theta_k))\,,$$

their sum is a martingale, and we reach

$$\mathbb{E}[e^{-\beta\,\theta_k}\varphi(\theta_k)] \geq \mathbb{E}[e^{-\beta\,\theta_{k+1}}\,\varphi(\theta_{k+1})] + \mathbb{E}\left[\int_{\theta_k}^{\theta_{k+1}} e^{-\beta\, t}U_p(C_t)\,\mathrm{d}t\right].$$

Iterating this procedure for all the indexes k we obtain

$$\varphi(x, y, \lambda) \geq \mathbb{E}\left[\int_0^{\infty} U_p(C_t)\,\mathrm{d}t\right],$$

since $\lim_{t\to+\infty} e^{-\beta\, t}\varphi(t) \geq 0$. □

Let us give an explicit supersolution to obtain a dominant to the value function.

Corollary 7.4.2 *Let γ be a positive constants such that $1 - \mu_M \leq \gamma \leq \frac{1}{1-\mu_L}$, and define the function $B : [0,1] \mapsto \mathbb{R}$ by*

$$B(p) := \frac{\beta - r\,p}{p} - \frac{(\alpha - r)^2}{2\,\sigma^2\,(1-p)}\,.$$

Then for $p \in (0, \overline{p})$, and $\overline{p} \in (0,1)$ the unique solution of $B(\overline{p}) = 0$ on the domain of B, we have the following upper bound for the value function solution to (7.3.2):

$$v(x, y, \lambda) \leq \frac{A^{p-1}(p)}{p}(x + \gamma\, y)^p\,, \quad \forall (x, y, \lambda) \in (\mathbb{R}^+)^2 \times [b, \infty)\,, \tag{7.4.2}$$

where $A(p) = \frac{p}{1-p} B(p)$.

Proof. We have to prove that $\varphi(x,y,\lambda) := \frac{A^{p-1}(p)}{p}(x+\gamma\, y)^p$ satisfies inequality (7.4.1). The second and third term in the minimum operator are greater than or equal to zero by definition of γ, while for the first one we have the following expression

$$\begin{aligned}\beta\,\varphi - \mathcal{A}\varphi - \mathcal{J}\varphi - \widetilde{U}_p(\varphi_x) =& A^{p-1}(p)\,(x+\gamma y)^p \Big[\frac{\beta - rp}{p} - \frac{(\alpha-r)^2}{2\sigma^2(1-p)} - \frac{1-p}{p}A(p) \\ &+ \frac{1}{2(1-p)}\Big(\frac{\sigma(1-p)\gamma y}{x+\gamma y} - \frac{\alpha - r}{\sigma}\Big)^2 \\ &- \frac{\lambda}{p}\int_{-1}^{\infty}\Big(\Big(\frac{x+\gamma y(1+\eta)}{x+\gamma y}\Big)^p - 1\Big)\nu(\mathrm{d}\eta)\Big]\end{aligned} \tag{7.4.3}$$

$$\geq A^{p-1}(p)\,(x+\gamma y)^p\Big[\frac{\beta-rp}{p} - \frac{(\alpha-r)^2}{2\sigma^2(1-p)} - \frac{1-p}{p}A(p)\Big] \tag{7.4.4}$$

$$\geq A^{p-1}(p)\,(x+\gamma y)^p\Big[B(p) - \frac{1-p}{p}A(p)\Big], \tag{7.4.5}$$

where (7.4.3) is obtained by rearranging the terms in the first member in (7.4.1), and (7.4.4) is obtained by the following upper bound on the integral

$$\begin{aligned}\int_{-1}^{\infty}\Big(\Big(\frac{x+\gamma y(1+\eta)}{x+\gamma y}\Big)^p - 1\Big)\nu(\mathrm{d}\eta) &= \int_{-1}^{\infty}\Big(1 + \frac{\gamma\, y}{x+\gamma y}\,\eta\Big)^p\nu(\mathrm{d}\eta) - 1 \\ &\leq \Bigg(1 + \frac{\gamma\, y}{x+\gamma y}\int_{-1}^{\infty}\eta\,\nu(\mathrm{d}\eta)\Bigg)^p - 1 = 0\,.\end{aligned}$$

This comes from the fat that $f : [-1,\infty) \mapsto \mathbb{R}$ defined as $f(x) = (1 + c\,x)^p$ is concave $\forall c, p \in [0,1]$, then we can apply Jensen's inequality, and we reach the final result by using the fact that the measure ν has negative expectation. Then, since (7.4.5) is equal to zero, we have the result as a consequence of Proposition 7.4.1. ☐

Remark 7.4.3 The upper bound for the value function given by (7.4.2) setting $\gamma = 1$ and $\lambda = 0$ is the one found in [89] with no transaction costs and no jump component, i.e. $\mu_M, \mu_L = 0$ and $\lambda \equiv 0$.

Let us show that the value function is continuous on its whole domain, i.e. $\forall(x,y,\lambda) \in \overline{\mathcal{S}}\times[b,+\infty)$, and therefore by the concavity property in Proposition 7.2.1 and the lower bound in Proposition 7.2.2 we will also obtain that the value function is increasing.

Proposition 7.4.4 *The value function v defined by (7.1.4) is continuous in the interior of its domain: $\mathcal{S}\times(b,+\infty)$.*

Proof. Continuity with respect to to (x, y) in $\mathcal{S}$ follows by concavity, see Proposition 7.2.1, and Corollary 7.4.2.

Let us prove that the value function is jointly continuous in the interior of its domain. Consider $(x, y, \lambda) \in \mathcal{S} \times (b, +\infty)$ and a positive constant γ such that $\gamma < \kappa/2$, and define the first exit time from the cube centered in (x, y, λ) with side length 2γ as

$$T^{\gamma} := \inf \left\{ t \geq 0 : \max\{|X_t - \gamma|, |Y_t - \gamma|, |\lambda_t - \gamma|\} \geq 0 \right\},$$

and similar definitions for the exit times from the intervals centered respectively in x, y, λ:

$$\begin{aligned} T^{X,\gamma} &:= \inf\{t \geq 0 : |X_t - \gamma| \geq 0\}, \\ T^{Y,\gamma} &:= \inf\{t \geq 0 : |Y_t - \gamma| \geq 0\}, \\ T^{\lambda,\gamma} &:= \inf\{t \geq 0 : |\lambda_t - \gamma| \geq 0\}. \end{aligned}$$

Notice that by definition of γ we have $\theta_1 \geq T^{\lambda,\gamma}$, where by θ_1 we denote the first arrival time for the Hawkes process, and moreover we have $T^{\gamma} = \min\{T^{X,\gamma}, T^{Y,\gamma}, T^{\lambda,\gamma}\}$.

We are going to prove that

$$v(X_{T^{\gamma}}, Y_{T^{\gamma}}, \lambda_{T^{\gamma}}) \to v(x, y, \lambda), \quad \text{a.s. for } \gamma \to 0,$$

where by $(X_{T^{\gamma}}, Y_{T^{\gamma}}, \lambda_{T^{\gamma}})$ we mean the process following the optimal strategy.

- Let us now consider the auxiliary process $(X^{(0,0,0)}, Y^{(0,0,0)}, \lambda)$ not considering transactions and consumption, and the auxiliary process $(X^{(0,0,0)}, \widetilde{Y}^{(0,0,0)}, \widetilde{\lambda})$ not considering any arrival from the jump process, no transactions and no consumption:

$$\begin{aligned} \mathrm{d}Y_t^{(0,0,0)} &= \alpha\, Y_t^{(0,0,0)}\, \mathrm{d}t + \sigma\, Y_t^{(0,0,0)}\, \mathrm{d}W_t + Y_{t^-}^{(0,0,0)} \int_{-1}^{+\infty} \eta\, \mathcal{N}(\mathrm{d}t, \mathrm{d}\eta), \\ \mathrm{d}X_t^{(0,0,0)} &= r\, X_t^{(0,0,0)}\, \mathrm{d}t, \\ \mathrm{d}\widetilde{Y}_t^{(0,0,0)} &= \alpha\, \widetilde{Y}_t^{(0,0,0)}\, \mathrm{d}t + \sigma\, \widetilde{Y}_t^{(0,0,0)}\, \mathrm{d}W_t, \\ \mathrm{d}\widetilde{\lambda}_t &= a\,(b - \widetilde{\lambda}_t)\, \mathrm{d}t, \end{aligned}$$

  notice that $X^{(0,0,0)}$ is deterministic and increasing in time, $\widetilde{\lambda}$ deterministic and decreasing in time, and that $(X^{(0,0,0)}, Y^{(0,0,0)})$ corresponds to the portfolio without transactions and consumption.

By the DPP we have

$$v(x,y,\lambda) = \sup_{(C,L,M)\in\mathcal{A}_{[0,T^\gamma]}(x,y)} \mathbb{E}\left[\int_0^{T^\gamma} e^{-\beta\,t}\,U_p(C_t)\,\mathrm{d}t + e^{-\beta\,T^\gamma}\,v(X_{T^\gamma},Y_{T^\gamma},\lambda_{T^\gamma})\right],$$

therefore, by considering the sub-optimal strategy consisting in no transactions and no consumption, we have $\mathbb{E}\left[v(X^{(0,0,0)}_{T^\gamma},Y^{(0,0,0)}_{T^\gamma},\lambda_{T^\gamma}) - v(x,y,\lambda)\right] \le 0$, moreover

$$\begin{aligned}
&\mathbb{E}\Big[v(X^{(0,0,0)}_{T^\gamma},Y^{(0,0,0)}_{T^\gamma},\lambda_{T^\gamma}) - v(x,y,\lambda)\Big] \\
&\qquad\qquad \ge \mathbb{E}\left[v(x,Y^{(0,0,0)}_{T^\gamma},\lambda_{T^\gamma}) - v(x,y,\lambda)\right] \\
&\qquad\qquad \ge \mathbb{E}\left[\Big(v(x,Y^{(0,0,0)}_{T^\gamma},(\lambda-b)e^{a\theta_1}-b+\kappa) - v(x,y,\lambda)\Big)\mathbb{1}_{T^\gamma=\theta_1}\right] \\
&\qquad\qquad\quad + \mathbb{E}\left[\Big(v(x,\widetilde{Y}^{(0,0,0)}_{T^\gamma},\lambda) - v(x,y,\lambda)\Big)\,\mathbb{1}_{T^\gamma>\theta_1}\right] \\
&\qquad\qquad \ge (v(x,0,\lambda+\kappa) - v(x,y,\lambda))\,\mathbb{P}(T^\gamma=\theta_1) \\
&\qquad\qquad\quad + \mathbb{E}\left[v(x,\widetilde{Y}^{(0,0,0)}_{T^\gamma},\lambda) - v(x,y,\lambda)\right],
\end{aligned}$$

where $\mathbb{E}[v(x,\widetilde{Y}^{(0,0,0)}_{T^\gamma},\lambda) - v(x,y,\lambda)]$ goes to zero for $\gamma\to 0$, since, being a geometric Brownian motion, $\widetilde{Y}^{(0,0,0)}$ has continuous paths, and by the continuity of v with respect to the second component. Moreover, also $\mathbb{P}(\theta_1 = T^\gamma)$ goes to zero for $\gamma \to 0$, since

$$\begin{aligned}
\mathbb{P}(\theta_1 = T^\gamma) &= \mathbb{P}\left(\theta_1 = \min\{T^{X,\gamma},T^{Y,\gamma},T^{\lambda,\gamma}\}\right) \\
&\le \mathbb{P}\left(\theta_1 \le T^{\widetilde{\lambda},\gamma}\right) \\
&\le \mathbb{P}\left(\theta_1 \le \frac{1}{a}\log\left(1+\frac{\gamma}{\lambda-b-\gamma}\right)\right) \\
&= 1-\left(1+\frac{\gamma}{\lambda-b-\gamma}\right)^{-\lambda/a},
\end{aligned}\tag{7.4.6}$$

since $T^{\lambda,\gamma} < T^{\widetilde{\lambda},\gamma}$, and the (7.4.6) tends to zero for $\gamma\to 0$.

- By the DPP we have

$$\begin{aligned}
&\mathbb{E}[v(X_{T^\gamma},Y_{T^\gamma},\lambda_{T^\gamma}) - v(x,y,\lambda)] \\
&\quad \le \mathbb{E}[(1-e^{-\beta\,T^\gamma})\,v(X_{T^\gamma},Y_{T^\gamma},\lambda_{T^\gamma})] \\
&\quad \le \mathbb{E}[(1-e^{-\beta\,T^\gamma})\,v(X_{T^\gamma},Y_{T^\gamma},\lambda_{T^\gamma})\mathbb{1}_{T^\gamma<\theta_1}]
\end{aligned}\tag{7.4.7}$$

$$
\begin{aligned}
&\quad + \mathbb{E}[(1-e^{-\beta\,\tau_1})\, v(X_{\tau_1}, Y_{\tau_1}, \lambda_{\tau_1}) \mathbb{1}_{T^\gamma \geq \theta_1}] \\
&\leq \mathbb{E}[(1-e^{-\beta\,T^\gamma})]\, v(x+y+2\,\gamma, x+y+2\,\gamma, \lambda-\gamma) \\
&\quad + \mathbb{E}[v(x+y+2\,\gamma, Y_{\tau_1}, \lambda-\gamma)\, \mathbb{1}_{T^\gamma \geq \theta_1}] \\
&\leq \mathbb{E}[(1-e^{-\beta\,T^\gamma})]\, v(x+y+2\,\gamma, x+y+2\,\gamma, \lambda-\gamma) + \mathbb{P}(\theta_1 \leq T^\gamma) \\
&\quad \times \left(\frac{A^{p-1}(p)}{p} \int_{-1}^{\infty} (x+y+2\,\gamma+\gamma\,(x+y+2\,\gamma)\,(1+\eta))^p \ \nu(\mathrm{d}\eta) \right)
\end{aligned} \tag{7.4.8}
$$

$$
\begin{aligned}
&\leq \mathbb{E}[(1-e^{-\beta\,T^\gamma})]\, v(x+y+2\,\gamma, x+y+2\,\gamma, \lambda-\gamma) + \mathbb{P}(\theta_1 \leq T^\gamma) \\
&\quad \times \left(\frac{A^{p-1}(p)}{p} \left((x+y+2\,\gamma)\,(1+\gamma) + \gamma\,(x+y+2\,\gamma) \int_{-1}^{\infty} \eta\, \nu(\mathrm{d}\eta) \right) \right)
\end{aligned} \tag{7.4.9}
$$

$$
\begin{aligned}
&\leq \left(\frac{\gamma}{\gamma+b-\lambda} \right)^{-\beta/a} v(x+y+2\,\gamma, x+y+2\,\gamma, \lambda-\gamma) \\
&\quad + \frac{A^{p-1}(p)}{p}\,(x+y+2\,\gamma)\,(1+\gamma)\, \mathbb{P}(\theta_1 \leq T^\gamma)\,,
\end{aligned} \tag{7.4.10}
$$

where inequality (7.4.8) is direct consequence of Corollary 7.4.2 and the independence of $v(x+y+2\,\gamma, x+y+2\,\gamma+\eta_1, \lambda-\gamma)$ by $\mathbb{1}_{T^\gamma \geq \theta_1}$, where η_1 is the size of the first jump that arrives at time θ_1, inequality (7.4.9) comes by Jensen's inequality, then (7.4.10) is obtained by the fact that $T^\gamma \leq T^{\lambda,\gamma} = T^{\widetilde{\lambda},\gamma} \wedge \theta_1$ implies that

$$
T^\gamma \ \leq \ T^{\widetilde{\lambda},\gamma} \ = \ \frac{1}{a} \log\left(1 + \frac{\gamma}{\lambda-b-\gamma}\right),
$$

and finally we have (7.4.10) since the Lévy measure has negative expectation.

Therefore, by the boundedness of v and by (7.4.6), the last expression tends to zero for $\gamma \to 0$. □

7.5 Viscosity solution

Lemma 7.5.1 *The value function defined by* (7.1.4) *is a viscosity supersolution to the system of variational inequalities* (7.3.2).

Proof. Let $(\overline{x}, \overline{y}, \overline{\lambda}) \in \overline{\mathcal{S}} \times [b, +\infty)$, and consider a test function $\varphi \in C^2$ such that $\varphi \leq v$ and $\varphi(\overline{x}, \overline{y}, \overline{\lambda}) = v(\overline{x}, \overline{y}, \overline{\lambda})$. Therefore, defining $B_\epsilon^+ := \{(x,y,\lambda) \in \overline{\mathcal{S}} \times [b, +\infty) : |x-\overline{x}| \leq \epsilon, |y-\overline{y}| \leq \epsilon, |\lambda-\overline{\lambda}| \leq \epsilon\}$, we clearly have that $(\overline{x}, \overline{y}, \overline{\lambda})$ minimizes $v-\varphi$

over B_ϵ^+. By Proposition 7.2.2 we have

$$\varphi(\overline{x},\overline{y},\overline{\lambda}) = v(\overline{x},\overline{y},\overline{\lambda}) \geq \frac{1}{p} C_*^{p-1} \left(x + (1-\mu_{M})\, y\right)^p .$$

Consider the strategy consisting in transferring the quantity $\delta < \epsilon \wedge \overline{x}$ from the risk-free asset to the risky asset, therefore we have

$$\varphi(\overline{x},\overline{y},\overline{\lambda}) = v(\overline{x},\overline{y},\overline{\lambda}) \geq v(\overline{x}-\delta, \overline{y}+\delta\,(1-\mu_{L}), \overline{\lambda}) ,$$

and also $\varphi(\overline{x},\overline{y},\overline{\lambda}) - \varphi(\overline{x}-\delta, \overline{y}+\delta\,(1-\mu_{L}), \overline{\lambda}) \geq 0$. Hence dividing by $||[-\delta,(1-\mu_{L})\,\delta]||$, and taking the limit $\delta \to 0$, we obtain

$$\nabla\varphi \cdot [-1;(1-\mu_{L})] \geq 0 .$$

A similar argument works also by transferring an amount δ from the asset to the stock to get

$$\nabla\varphi \cdot [(1-\mu_{M});-1] \geq 0 ,$$

and we are left with proving that

$$\beta\,\varphi - A\varphi - \mathcal{J}\varphi - \widetilde{U}_p(\varphi_x) \geq 0 .$$

Now we have to prove the inequality for the first member in (7.4.1), i.e.

$$\beta\,\varphi(\overline{x},\overline{y},\overline{\lambda}) - A\varphi(\overline{x},\overline{y},\overline{\lambda}) - \mathcal{J}\varphi(\overline{x},\overline{y},\overline{\lambda}) - \widetilde{U}_p(\varphi_x(\overline{x},\overline{y},\overline{\lambda})) \geq 0 .$$

Let h be a stopping time, θ the first arrival in the Hawkes process, define $\gamma_\epsilon := t_\epsilon \wedge h \wedge \theta$, and consider an admissible strategy $(C,L,M) \in \mathcal{A}(\overline{x},\overline{y})$ such that $L = M \equiv 0$ and $C_t = c$ for $t \in [0,t_\epsilon)$, then by the DPP we have

$$\begin{aligned}
\varphi(\overline{x},\overline{y},\overline{\lambda}) = v(\overline{x},\overline{y},\overline{\lambda}) \geq & \mathbb{E}\bigg[\int_0^{\gamma_\epsilon} e^{-\beta\, s} U_p(c)\,\mathrm{d}s + v\Big(X^{(c,0,0)}_{h\wedge t_\epsilon}, Y^{(c,0,0)}_{h\wedge t_\epsilon}, \lambda_{h\wedge t_\epsilon}\Big)\,\mathbb{1}_{h\wedge t_\epsilon<\theta} \\
& + v\Big(X^{(c,0,0)}_{\theta^-}, Y^{(c,0,0)}_{\theta^-}(1+\eta), \lambda_{\theta^-}+\kappa\Big)\,\mathbb{1}_{\theta<h\wedge t_\epsilon}\bigg] \\
\geq & \mathbb{E}\bigg[\int_0^{\gamma_\epsilon} e^{-\beta\, s} U_p(c)\,\mathrm{d}s + \varphi\Big(X^{(c,0,0)}_{h\wedge t_\epsilon}, Y^{(c,0,0)}_{h\wedge t_\epsilon}, \lambda_{h\wedge t_\epsilon}\Big)\,\mathbb{1}_{h\wedge t_\epsilon<\theta} \\
& + \varphi\Big(X^{(c,0,0)}_{\theta^-}, Y^{(c,0,0)}_{\theta^-}(1+\eta), \lambda_{\theta^-}+\kappa\Big)\,\mathbb{1}_{\theta<h\wedge t_\epsilon}\bigg], \qquad (7.5.1)
\end{aligned}$$

where η is the size of the jump arrive at time θ and $(X^{(C,L,M)}, Y^{(C,L,M)})$ is the process given by equations (7.1.2) and (7.1.3) applying an admissible strategy (C, L, M). Apply Itô formula to $e^{-\beta t}\varphi(X_t^{(c,0,0)}, Y_t^{(c,0,0)}, \lambda_t)$ for times $t \in [0, \gamma_\epsilon)$, and take the expectation to get

$$\begin{aligned}
&\mathbb{E}[e^{-\beta\gamma_\epsilon}\varphi(X_{\gamma_\epsilon}^{(c,0,0)}, Y_{\gamma_\epsilon}^{(c,0,0)}, \lambda_{\gamma_\epsilon})] \\
&\qquad = \varphi(\overline{x}, \overline{y}, \overline{\lambda}) + \mathbb{E}\left[\int_0^{\gamma_\epsilon} e^{-\beta t}\,(-\beta\,\varphi + A\varphi - c\,\varphi_x)(X_t^{(c,0,0)}, Y_t^{(c,0,0)}, \lambda_t)\,\mathrm{d}t\right].
\end{aligned}$$

Then combining it with (7.5.1), and noticing that $\mathbb{E}[e^{-\beta t}\varphi] \le \mathbb{E}[\varphi]$, we obtain

$$\begin{aligned}
0 \ge & \mathbb{E}\left[\int_0^{\gamma_\epsilon} e^{-\beta t}\,(-\beta\,\varphi + A\varphi - c\,\varphi_x)(X_t^{(c,0,0)}, Y_t^{(c,0,0)}, \lambda_t) + U_p(c)\,\mathrm{d}t\right] \\
& + \mathbb{E}\left[\varphi\big(X_{\theta^-}^{(c,0,0)}, Y_{\theta^-}^{(c,0,0)}(1+\eta), \lambda_{\theta^-} + \kappa\big)\,\mathbb{1}_{\theta < h\wedge t_\epsilon}\right] \\
\ge & \mathbb{E}\left[\int_0^{\gamma_\epsilon} e^{-\beta t}\,(-\beta\,\varphi + A\varphi + \mathcal{J}\varphi - c\,\varphi_x)(X_t^{(c,0,0)}, Y_t^{(c,0,0)}, \lambda_t) + U_p(c)\,\mathrm{d}t\right],
\end{aligned}$$

whose integrand part is bounded by definition of t_ϵ, therefore letting $\epsilon \to 0$ and taking the supremum over $c > 0$, we obtain the supersolution property for the value function. □

Lemma 7.5.2 *The value function defined by (7.1.4) is a viscosity subsolution to the system of variational inequalities (7.3.2).*

Proof. To prove the subsolution property we proceed by contradiction, i.e. we assume that there exist a constant $\epsilon > 0$, a point $(\overline{x}, \overline{y}, \overline{\lambda}) \in \overline{\mathcal{S}} \times [b, +\infty)$, a C^2-function φ such that $(\varphi - v)(\overline{x}, \overline{y}, \overline{\lambda}) = 0$ and $\varphi \ge v$ on a neighborhood $B_\epsilon(\overline{x}, \overline{y}, \overline{\lambda}) := \{(x, y, \lambda) \in \overline{\mathcal{S}} \times [b, +\infty) : |x - \overline{x}| \le \epsilon, |y - \overline{y}| \le \epsilon, |\lambda - \overline{\lambda}| \le \epsilon\}$, and $\zeta > 0$ such that for all $(x, y, \lambda) \in B_\epsilon(\overline{x}, \overline{y}, \overline{\lambda})$ we have

$$\beta\,\varphi(x,y,\lambda) - A\varphi(x,y,\lambda) - \mathcal{J}\varphi(x,y,\lambda) - \widetilde{U}_p(\varphi_x(x,y,\lambda)) > \zeta\,, \tag{7.5.2}$$

$$-(1-\mu_M)\,\partial_x\varphi(x,y,\lambda) + \varphi_y(x,y,\lambda) > \zeta\,, \tag{7.5.3}$$

$$\varphi_x(x,y,\lambda) - (1-\mu_L)\,\varphi_y(x,y,\lambda) > \zeta\,. \tag{7.5.4}$$

Consider the exit time form the ball

$$t_\epsilon := \inf\{t \ge 0 : (X_t, Y_t, \lambda_t) \notin B_\epsilon(\overline{x}, \overline{y}, \overline{\lambda})\}\,,$$

and define the time $\gamma_\epsilon := t_\epsilon \wedge \tau$, for $\tau \in [0,\infty)$. Then apply Itô formula to $e^{-\beta t}\varphi(X_t, Y_t, \lambda_t)$ for $t \in [0, t_\epsilon)$

$$
\begin{aligned}
&e^{-\beta\gamma_\epsilon}\varphi(\gamma_\epsilon)\\
&\quad= \varphi(\overline{x},\overline{y},\overline{\lambda}) + \int_0^{\gamma_\epsilon} e^{-\beta t}(-\beta\varphi(t) + A\varphi(t) - C_t\,\varphi_x(t))\,\mathrm{d}t + \int_0^{\gamma_\epsilon} e^{-\beta t}\,\sigma\,Y_t\,\varphi_y(t)\,\mathrm{d}W_t\\
&\qquad + \sum_{0<t<\gamma_\epsilon} e^{-\beta t}\Big((\lambda_t - \lambda_{t^-})\,\partial_\lambda\varphi(t) + (Y_t - Y_{t^-})\,\partial_y\varphi(t)\Big)\\
&\qquad + \int_0^{\gamma_\epsilon} e^{-\beta t}\,((1-\mu_M)\,\varphi_y(t) - \varphi_x(t))\mathrm{d}L_t + \int_0^{\gamma_\epsilon} e^{-\beta t}((1-\mu_L)\varphi_x(t) - \varphi_y(t))\mathrm{d}M_t\,,
\end{aligned}
$$

and take the expectation to get

$$
\begin{aligned}
\varphi(\overline{x},\overline{y},\overline{\lambda}) \geq& \mathbb{E}[e^{-\beta\gamma_\epsilon}\varphi(\gamma_\epsilon)] + \mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\big(\beta\varphi(t) - A\varphi(t) - \mathcal{J}\varphi(t) + C_t\,\varphi_x(t)\big)\,\mathrm{d}t\Big]\\
&+ \zeta\,\mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\,(\mathrm{d}L_t + \mathrm{d}M_t)\Big]\\
=&\mathbb{E}[e^{-\beta\gamma_\epsilon}\varphi(\gamma_\epsilon)] + \mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\big(\beta\varphi(t) - A\varphi(t) - \mathcal{J}\varphi(t) - \widetilde{U}_p(\varphi_x(t))\big)\,\mathrm{d}t\Big]\\
&+ \zeta\,\mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\,(\mathrm{d}L_t + \mathrm{d}M_t)\Big] + \mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\big(\widetilde{U}_p(\varphi_x(t)) + C_t\,\varphi_x(t)\big)\Big]\\
\geq&\mathbb{E}[e^{-\beta\gamma_\epsilon}\varphi(\gamma_\epsilon)] + \mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\,U_p(C_t)\,\mathrm{d}t\Big] + f(t), \qquad (7.5.5)
\end{aligned}
$$

where

$$
f(t) = \inf_{(C,L,M)\in\mathcal{A}(\overline{x},\overline{y})}\Big\{\zeta\,e^{-\beta\gamma_\epsilon}\,\mathbb{E}[\gamma_\epsilon + L_{\gamma_\epsilon} + M_{\gamma_\epsilon}] + \mathbb{E}\Big[\int_0^{\gamma_\epsilon} e^{-\beta t}\big(\widetilde{U}_p(\varphi_x(t)) - (U_p(C_t) - C_t\,\varphi_x(t))\big)\,\mathrm{d}t\Big]\Big\}.
$$

Therefore taking the supremum over the admissible strategies in the right hand side of (7.5.5) we have

$$
\varphi(\overline{x},\overline{y},\overline{\lambda}) \;\geq\; f(t) + v(\overline{x},\overline{y},\overline{\lambda}) \;\geq\; f(t) + \varphi(\overline{x},\overline{y},\overline{\lambda})
$$

hence, since for t sufficiently small $f(t) > 0$, we obtain our contradiction. ☐

Theorem 7.5.1 (Viscosity solution) *The value function v given by* (7.1.4) *is the viscosity solution to the system of variational inequalities* (7.3.2)*.*

Proof. This theorem is a direct consequence of Lemma 7.5.1 and Lemma 7.5.2. ☐

Chapter 8

Comparative statics

8.1 Reduction of the dimension of the problem

We reduce the dimension of the problem by one variable. Define

$$u(z,\lambda) = v(1-z, z, \lambda)\,, \quad \text{for } z \in [0,1], \lambda \geq b\,,$$

then by the homotheticity property in Proposition 7.2.1, for all $(x,y,\lambda) \in \overline{S} \setminus \{0,0\} \times [b,\infty)$ we have

$$v(x,y,\lambda) = (x+y)^p \, u\Big(\frac{y}{x+y}, \lambda\Big). \tag{8.1.1}$$

For $z \in [0,1]$ and $\lambda \in [b,\infty)$, we define

$$\begin{aligned}
d_1(z) &= r + (\alpha - r)\, z - \frac{1}{2}\,\sigma^2\,(1-p)\,z^2 \\
d_2(z) &= (\alpha - r)\, z\,(1-z) - \sigma^2\,(1-p)\,z^2\,(1-z) \\
d_3(z) &= \frac{1}{2}\,\sigma^2\, z^2\,(1-z)^2 \\
d_4(z) &= \frac{1}{\mu_M}\,(1 - \mu_M\, z) \\
d_5(z) &= \frac{1}{\mu_L}\,(1 - \mu_L\,(1-z)) \\
\widetilde{\mathcal{J}}\psi(z,\lambda) &= \lambda \int_{-1}^{\infty} \left[(1+z\,\eta)^p\,\psi\left(\frac{z\,(1+\eta)}{z\,\eta+1}, \lambda+\kappa\right) - \psi(z,\lambda)\right] \nu(\mathrm{d}\eta)\,.
\end{aligned} \tag{8.1.2}$$

Proposition 8.1.1 *The value function v is solution to the HJB equation (7.3.2) if and only if u, given by (8.1.1), is solution to the two-dimensional second-order integral* *differential equation*

$$\min\{\beta\,\psi(z,\lambda) - d_1(z)\,p\,\psi(z,\lambda) - d_2(z)\,\partial_z\psi(z,\lambda) - d_3(z)\,\partial^2_{zz}\psi(z,\lambda) \tag{8.1.3}$$

$$-a\,(b-\lambda)\,\partial_\lambda\psi(z,\lambda) - \widetilde{\mathcal{J}}\psi(z,\lambda) - \widetilde{U}_p(p\,\psi(z,\lambda) - z\,\partial_z\psi(z,\lambda)),$$

$$p\,\psi(z,\lambda) + d_4(z)\,\partial_z\psi(z,\lambda),\; p\,\psi(z,\lambda) - d_5(z)\,\partial_z\psi(z,\lambda)\} = 0\,. \tag{8.1.4}$$

Proof. Let us start proving that the fact that the value function v satisfies equation (7.3.2) implies that u given by (8.1.1) satisfies (8.1.4). Consider the integral term in (7.3.2), by definition of (8.1.1) the integral term results in

$$\mathcal{J}v(x,y,\lambda) = \lambda\int_{-1}^{\infty}\Big((x+y\,(1+\eta))^p\,u\Big(\frac{y\,(1+\eta)}{x+y\,(1+\eta)},\lambda\Big) - (x+y)^p\,u\Big(\frac{y}{x+y},\lambda\Big)\Big)\,\nu(\mathrm{d}\eta)\,,$$

then consider the change of variable $\frac{y}{x+y} = z$, implying $\frac{x}{x+y} = 1-z$ and also

$$\frac{y\,(1+\eta)}{x+y\,(1+\eta)} = \frac{z\,(1+\eta)}{1+z\,\eta}\,, \qquad (x+y\,(1+\eta))^p = (x+y)^p\,(1+z\,\eta)^p\,,$$

then

$$\mathcal{J}v(x,y,\lambda) = (x+y)^p\,\widetilde{\mathcal{J}}u(z,\lambda)\,.$$

Proceed in the same way for all the terms in (7.3.2), then equation (8.1.4) satisfied by u comes by direct computation. For the other implication, simply start from (8.1.4) instead of (7.3.2). □

8.2 Numerical results

Let us start by a *penalty* approximation of the HJB (8.1.3) obtained after the convenient change of variable

$$a\,(b-\lambda)\,\partial_\lambda\psi^m + \mathscr{L}\psi^m + \mathcal{J}^m\psi^m + K\,\{(\mathcal{L}\psi^m)^+ + (\mathcal{M}\psi^m)^+\} = 0\,, \tag{8.2.1}$$

for $\lambda > b$, $m \in \mathbb{N}$, and where K is the penalty parameter big enough in order to penalize the solutions that are not satisfying the constraints of the original HJB

(8.1.4). Furthermore, we defined the operators as follows

$$\begin{aligned}\mathscr{L}\psi^m =&\left(-\frac{\beta}{\gamma}+r+(\alpha-r)\,z-\frac{1}{2}\sigma^2\,(1-\gamma)\,z^2\right)\gamma\,\psi^m+\frac{1-\gamma}{\gamma}\left(\gamma\,\psi^m-z\,\partial_z\psi^m\right)^{\frac{\gamma}{\gamma-1}}\\&+\left((\alpha-r)\,z\,(1-z)-\sigma^2\,(1-\gamma)\,z^2\,(1-z)\right)\partial_z\psi^m+\frac{1}{2}\,\sigma^2\,z^2\,(1-z)^2\,\partial^2_{zz}\psi^m\,,\\\mathcal{J}^m\psi^m =&\lambda\int_{-1}^{\infty}\left[(1+z\,\eta)^\gamma\,\psi^{m-1}\left(\frac{z\,(1+\eta)}{z\,\eta+1},\lambda+\kappa\right)-\psi^m(z,\lambda)-z\,\eta\,\partial_z\psi^m(z,\lambda)\right]\nu(\mathrm{d}\eta)\,,\end{aligned}\tag{8.2.2}$$

$$\begin{aligned}\mathcal{M}\psi^m =&(-1+\mu_M\,z)\,\partial_z\psi^m-\mu_M\,\gamma\,\psi^m\,,\\\mathcal{L}\psi^m =&(1-\mu_L\,(1-z))\,\partial_z\psi^m-\mu_L\,\gamma\,\psi^m\,.\end{aligned}$$

The solutions ψ^m of (8.2.1) are defined interactively, starting at $m=0$ where the integral operator is

$$\mathcal{J}^0\psi^0=\lambda\int_{-1}^{\infty}\left[(1+z\,\eta)^\gamma\,\psi^0\left(\frac{z\,(1+\eta)}{z\,\eta+1},\lambda\right)-\psi^0(z,\lambda)-z\,\eta\,\partial_z\psi^0(z,\lambda)\right]\nu(\mathrm{d}\eta)\,,$$

instead the one in (8.2.2). Notice that we expressed the integral operator (8.2.2) slightly differently in order to be able to consider also jump measures with non-null expectation while retaining the assumption of compensated jumps.

We remark the fact that the penalty equation (8.2.1) holds for $\lambda>b$, while we have the initial condition for $\lambda=b$ solution of the same Penalty HJB but with the difference that there is no exponential time-decay in the intensity of the jump process.

To avoid numerical oscillations, we will consider the following change of variables:

$$W(z,\lambda)=\frac{\log(\gamma\,\psi)}{\gamma}\,,$$

and the corresponding equation is

$$a\,(b-\lambda)\,\partial_\lambda W^m+\mathscr{L}_1W^m+\mathcal{J}_1^mW^m+K\left\{(\mathcal{L}_1W^m)^++(\mathcal{M}_1W^m)^+\right\}=0\,,\tag{8.2.3}$$

for $\lambda>b$, where, letting $\overline{\nu}=\lambda\int_0^\infty\eta\,\nu(\mathrm{d}\eta)$ be the compensator of the jumps, we defined the operators as follows

$$\begin{aligned}\mathscr{L}_1W^m =&\left(-\frac{\beta+\lambda}{\gamma}+r+(\alpha-\overline{\nu}-r)\,z-\frac{1}{2}\sigma^2\,(1-\gamma)\,z^2\right)+\frac{1-\gamma}{\gamma}\left(1-z\,\partial_zW^m\right)^{\frac{\gamma}{\gamma-1}}\\&+\left((\alpha-\overline{\nu}-r)\,z\,(1-z)-\sigma^2\,(1-\gamma)\,z^2\,(1-z)\right)\partial_zW^m\\&+\frac{1}{2}\,\sigma^2\,z^2\,(1-z)^2\left(\partial^2_{zz}W^m+\gamma\,(\partial_zW^m)^2\right)\,,\end{aligned}$$

$$
\begin{aligned}
\mathcal{J}_1^m W^m =&\frac{\lambda}{\gamma}\, e^{-\gamma\, W^m} \int_{-1}^{\infty} (1+z\,\eta)^\gamma \, \exp\left(\gamma\, W^{m-1}\left(\frac{z\,(1+\eta)}{z\,\eta+1}, \lambda+\kappa\right)\right)\, \nu(\mathrm{d}\eta)\,,\\
\mathcal{M}_1 W^m =&(-1+\mu_M\, z)\, \partial_z W^m - \mu_M\, W^m\,,\\
\mathcal{L}_1 W^m =&(1-\mu_L\,(1-z))\, \partial_z W^m - \mu_L\, W^m\,,
\end{aligned}
$$

for $m > 1$, and for $m = 0$ we have the following integral operator

$$
\mathcal{J}_1^0 W^0 = \frac{\lambda}{\gamma}\, e^{-\gamma\, W^0} \int_{-1}^{\infty} (1+z\,\eta)^\gamma \, \exp\left(\gamma\, W^0\left(\frac{z\,(1+\eta)}{z\,\eta+1}, \lambda\right)\right)\, \nu(\mathrm{d}\eta)\,.
$$

8.2.1 Discretization

We perform the following discretizations:

- Truncation of the λ interval with uniform discretization

$$
b = \lambda_0 \quad < \quad \lambda_0 + \Delta\lambda = \lambda_1 \quad < \quad \dots \quad < \quad \lambda_N = \lambda_{\max}\,,
$$

  where $\Delta\lambda = (\lambda_{\max} - b)/N$.

- Space discretization

$$
0 = z_0 \quad < \quad z_0 + h = z_1 \quad < \quad \dots \quad < \quad z_L = 1\,,
$$

  where the discretization step $h = 1/L$.

- Jump size truncation and discretization.

- Finite difference approximation for $\partial_\lambda W$, $\partial_z W$, $\partial^2_{zz} W$.

To simplify the notation, let us consider again the case $m = 0$. Through the Implicit Euler λ-discretization we have

$$
\begin{aligned}
-\, a\,(b-\lambda_{n+1})\,\frac{W^{n+1}-W^n}{\Delta\lambda} - \mathscr{L}_1 W^{n+1} - \frac{1-\gamma}{\gamma}\left(1 - z\,\partial_z W^{n+1}\right)^{\frac{\gamma}{\gamma-1}} =&\\
K\left\{(\mathcal{L}_1 W^{n+1})^+ + (\mathcal{M}_1 W^{n+1})^+\right\}&\,,
\end{aligned}
$$

and collecting the unknown in the left-hand side we have

$$
\begin{aligned}
\left\{I + \frac{\Delta\lambda}{a\,(b-\lambda_{n+1})}\left(\mathscr{L}_{\text{lin}}\cdot + \frac{1-\gamma}{\gamma}\,(1 - z\,\partial_z\cdot)^{\frac{\gamma}{\gamma-1}} + K\left\{\mathcal{L}_1^*\cdot + \mathcal{M}_1^*\cdot\right\}\right)\right\} W^{n+1,l+1} =&\\
W^n - \frac{\Delta\lambda}{a\,(b-\lambda_{n+1})}\,\mathscr{L}_{\text{non-lin}} W^{n+1,l}\,,& \quad (8.2.4)
\end{aligned}
$$

where we splitted the operator $\mathscr{L}$ in linear and non-linear part as follows

$$\begin{aligned}\mathscr{L}_{\text{lin}}W =&((\alpha-r)\,z\,(1-z)-\sigma^2\,(1-\gamma)\,z^2\,(1-z))\,\partial_z W+\frac{1}{2}\,\sigma^2\,z^2\,(1-z)^2\,\partial^2_{zz}W\,,\\ \mathscr{L}_{\text{non-lin}}W =&-\frac{\beta+\lambda\,\widetilde{\nu}}{\gamma}+r+(\alpha-r)\,z-\frac{1}{2}\sigma^2\,(1-\gamma)\,z^2\\ &+\frac{1}{2}\,\sigma^2\,z^2\,(1-z)^2\gamma\,(\partial_z W)^2\\ &+\frac{\lambda}{\gamma}\,e^{-\gamma\,W}\int_{-1}^{\eta_{\max}}(1+z\,\eta)^\gamma\,\exp\left(W\left(\frac{z\,(1+\eta)}{z\,\eta+1},\lambda\right)\right)\,\nu(\mathrm{d}\eta)\,,\end{aligned}\tag{8.2.5}$$

and

$$\widetilde{\nu}=\int_{-1}^{\eta_{\max}}\nu(\mathrm{d}\eta)\,.$$

For the case $m>0$, the discretized integral operator (8.2.5) would be substituted by

$$\frac{\lambda}{\gamma}\,e^{-\gamma\,W^m}\int_{-1}^{\eta_{\max}}(1+z\,\eta)^\gamma\,\exp\left(W^{m-1}\left(\frac{z\,(1+\eta)}{z\,\eta+1},\lambda+\kappa\right)\right)\,\nu(\mathrm{d}\eta)\,.$$

8.3 Sensitivity Analysis

In this section we will start by fixing the following values for the parameters:

- $a=10$, $b=0.25$, $\kappa=0.5$, $\lambda_{\max}=10$, for the intensity of the jumps,
- $\alpha=0.12$, $\sigma=0.4$, $r=0$, $\beta=0.09$, for drift, volatility, risk-free interest rate and discount factor,
- $\gamma=0.4$, $\mu_M=0.01$, $\mu_L=0.01$, the utility function parameter and the transaction costs.
- The density on the size of the jumps in the risky asset:

$$\nu(\eta)=(\eta+1)^{\eta-1}\,\frac{e^{-(\eta+1)/\theta}}{\theta^k\,\Gamma(k)}\,,\tag{8.3.1}$$

  i.e. a gamma density function on $[-1,\infty)$ with mean value $k\,\theta-1$, which a priori has non-null expectation.

- $\theta\in\{0.4,\,0.5,\,0.75\}$, $k=2$, for the parameters of the gamma distribution for the sizes of the jumps in the risky asset, see (8.3.1).

In the case in which the expectation of the jump sizes is positive, but not too large, we have a weakened "buy region shrink" and "sell region expansion", see Figure 8.1

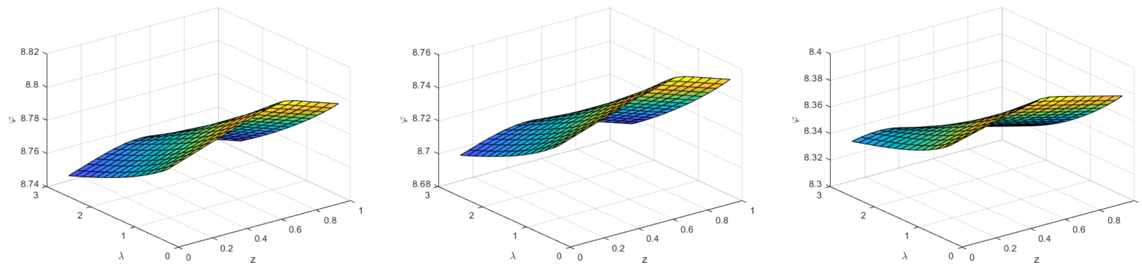

Figure 8.1: From the left to the right we represented the value functions surfaces in the cases in which the expectations of the jumps size are -0.2, 0 and 0.5.

for a representation of the surfaces corresponding to the value function for different expected value of the jumps, while see Figure 8.2 for a representation of the lines delimiting the sell and buy regions. Despite the fact that the jumps have positive expectation we have similar effects as in the previous cases and we still have that the value function is decreasing with respect to λ. This is mainly due to the fact that the utility function is concave and to the fact that the expectation on the jumps is only slightly positive.

In Figure 8.2 we compared the curves delimiting the buy/sell regions in case of self-exciting jumps and in the case of jumps with constant intensity. We plotted these graphs displaying the dependence with respect to the intensity of the jumps, but we remark the fact that, while in the Hawkes framework the intensity of the jumps λ is a true variable which varies in time, for Poisson jumps their intensity is only a fixed parameter.

Comparing the Hawkes-NTR and the Poisson-NTR one may notice that in the Hawkes case the NTR expands as the intensity increases, while for the Poisson case the NTR is almost constant in width. Moreover, looking at Figure 8.2, the buy region shrinks slower (and the sell region expands slower) than the buy (sell resp.) region in the Poisson framework. This is mainly due because we are considering a high speed of mean reversion coefficient $a = 10$.

8.4 Financial interpretation of the Hawkes jumps

This section is devoted to a comparison of the Hawkes-jump diffusion framework for a Merton type model and transaction costs with the the usual Poisson framework.

In the usual Poisson framework, when a positive (respectively negative) jump occurs the proportion $\pi^\star$ of the wealth invested in the risky asset increases (respectively

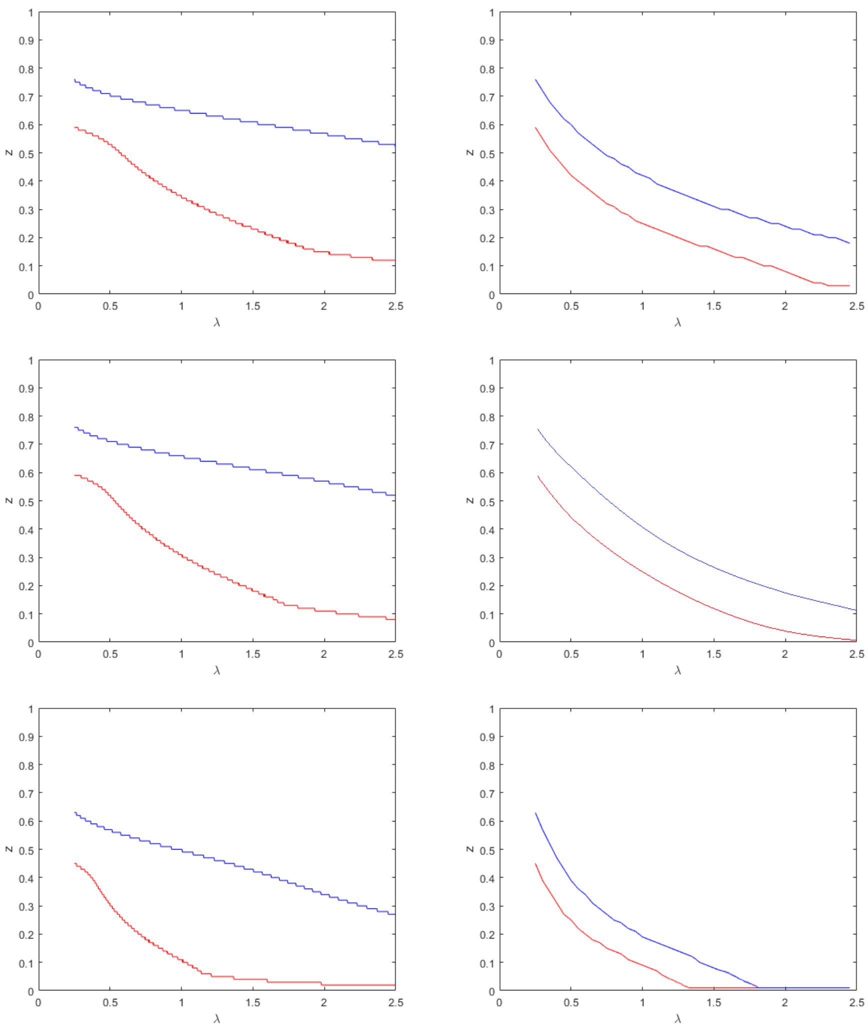

Figure 8.2: Buy/sell curves for a Merton-Hawkes and a Merton-Poisson framework. From the top down the expectation of the jump size is fixed to be -0.2, 0, 0.5 ($\theta = 0.4$, 0.5, 0.75, respectively, see equation (8.3.1)). The buy region is below the red curve, the sell region is above the blue curve and the no transaction region is between the two curves. In the left graphs we assumed the self-exciting component to be $\kappa = 0.5$, while in the right graphs we considered the case with intensity λ constant in time.

decreases) in order to keep this proportion constant. This result due to Merton [89] was extended in several papers, e.g. to cite some of them [4], [44], [58], [81], [109], and could be interpreted as a self-regulating rule for financial markets, since investors will sell (respectively buy) after a positive (respectively negative) fluctuation.

In contrast, in the Hawkes framework, both positive and negative jumps contribute to increase the intensity of future jumps. This sudden positive change in the intensity has an impact that is twofold on the optimal Merton proportion $\pi^\star$. First, increasing the frequency the variance of the risky asset is magnified and therefore the proportion $\pi^\star$ decreases, this phenomenon is outlined by Aït-Sahalia and Hurd [5] as a flight-to-quality due to risk aversion. Nevertheless the intensity has also a second potential effect related to the mean level of jumps affecting the mean level of stock return. Then, if jump size is non positive in average both previous effects are going in the same direction reducing the Merton proportion $\pi^\star$ during clustering periods. However, an ambiguous effect could rise if the jump size is positive in average. The financial explanation is that if clusters of positive jumps are allowed, the excess of positive return could compensate the excess of variance. In this part of the thesis, we focused mainly on the first effect due to the Hawkes process, since the jumps observed in financial markets are often negative. We then conclude that the optimal policy after a positive jump is always to sell the risky asset, as in the Poissonian case, but the quantity sold is larger, see Figure (6.8). In contrast, after a negative jump the optimal policy could be also to sell it, in contrast with the Merton rule, see Figure (8.3). This asymmetry is a by product of cluster structure of jumps combined with risk aversion.

We now turn on more general consequences of Hawkes framework showing that this feature could be seen as a self-fulfilling assumption joint with negative jumps. In order to do that, we relax the hypothesis of fully competitive markets since the assumption that investors can purchase and sell arbitrary large amounts of the risky asset at a fixed price per share is unrealistic due to liquidity constraints.

We consider now that a large part of agents are rational and consider the previous Hawkes framework, where a large investor decides to buy a large amount of risky asset (for exogenous reasons). The main effect is that the large investor pays a price per share larger than the one observed when she decided to buy, see for instance Kuhn and Stroh [79]. That is, a positive jump will appear, but there is no problem of lack of counterparties and then the decision of the large investor could be absorbed without other long term effects.

The situation is completely different if the same large investor wants to sell (always for exogenous reasons) since the optimal policy of the other agents is also to sell since they foresee an increasing frequency of future jumps, that is the main difference with the usual Poisson paradigm. There is then a lack of counterparties that is generally covered by opportunistic liquidity providers which are selling the excess of shares in the future - diluting the effect over large times. At the same time this time dilation could explain the existence of clusters. As a consequence jump clustering, the like as Hawkes type jumps, could be rise as a self-fulfilling prophecy.

This aspect could also participate to explain the observed volatility asymmetry with the leverage and volatility feedback effects. The strong dependency of the future volatility with respect to past negative jumps, compared to positive ones, detailed by Patton and Sheppard [99] could be explained by the previous mechanism.

Moreover, the asymmetry of the reaction of other agents to an external shock is clearly more important in the negative direction than in the positive one. In this way, the fact that the jumps are negative on average could be seen as a consequence of the clustering framework and, at the end, of the risk aversion.

Relaxing the hypothesis of fully competitive markets in order to study the impact of phenomena just detailed, we introduce proportional costs to buy and sell risky assets, that is a parsimonious way to take into account liquidity issues.

The main novelty of the presence of clustering had a double effect. First, we saw that the optimal buying and selling line are decreasing with lambda. But at the same time, the selling line never touches zero, forcing investors to keep a strictly positive proportion of wealth invested in the risky asset, even when the frequency λ goes to infinity, in contrast to the case without costs. More surprisingly, the clustering effect has another side, since after each jump the intensity λ sharply increases, but in absence of jumps the intensity smoothly decreases. This downhill has the effect that the buy line gradually turns anti-clockwise pushing investor to buy progressively the risky assets. As a consequence between two clusters, risky asset exhibits a positive excess of return due to the fact that agents progressively increase their exposure on the risky asset until the next cluster.

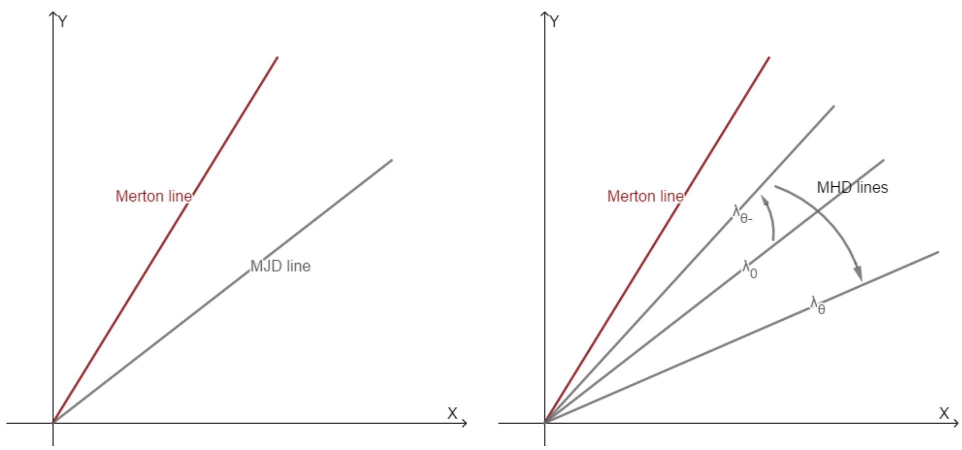

Figure 8.3: In the left figure the Merton Optimal allocation line (Merton line in short) is compared to a the Poisson Jump Diffusion optimal allocation line (MJD line in short), and shows that adding jumps in the model it becomes optimal to place a smaller wealth proportion in the risky asset , i.e. introducing jumps leads to a similar effect as increasing volatility. In the right figure the Merton line is compared to the Hawkes Jump Diffusion optimal allocation lines (MHD line in short) for different times. Here the parameters a, κ are assumed to be strictly positive and $\lambda_0 > b$. We represented the starting MHD line at time $t = 0$, the MHD line at the instant before of a jump arrival $t = \theta_1^-$ (which corresponds in continuously increasing the proportion of wealth invested in the risky asset for the period $(0, \theta_1)$), and the MHD line at the first jump arrival time $t = \theta_1$. From the figure it is clear that after a jump occurs the proportion of wealth invested in the risky asset increases dramatically.

Part V

Closed formula for options linked to target volatility strategies

There's no free lunch.

Sir John Templeton (1912-2008)

Introduction

In the aftermath of the financial markets, risk management solutions became more and more important for institutional and retail investors. The low interest rate environment forced practitioners to develop innovative and more efficient techniques to proficiently handling clients' portfolios risk budgets. To this aim, one of the most successful strategies, introduced both in multi asset portfolios and within structured products offering at least partial capital protection, is the so-called target volatility strategy. This concept shifts dynamically between risk-free and risky assets to generate a portfolio with a stable risk level independent of market volatilities. Such an approach assumes that market volatilities are a good indicator for asset allocation decisions. This method works well in rising markets with low volatility as well as in falling markets with higher volatilities. Practitioners often compare this solution with a constant portfolio protection insurance (CPPI) strategy, also dynamically allocating funds between risky and risk-free assets, see [56]. Nevertheless, CPPI also aims at achieving capital protection, see e.g. [66] for an overview on risk overlay portfolio strategies.

In recent years, dynamic asset allocation process as Target Volatility Strategies or CPPI Strategies were used as underlying of options and a series of academic papers studied option theory when the underlying of the derivative follows a certain trading rule shifting between risky and risk-free assets. Most of them took a numeric approach to determine option prices or hedging parameters. We refer to, e.g. Albeverio et al. [6, 7], Jawaid [69, 70], Zakamulin [117], to what concerns the analysis of Target Volatility Strategies in different market models, while Zagst et al. [56] focused on option on a CPPI, also developing a closed-end formula of CPPI options in a Black-Scholes environment.

The analysis of this part of the thesis is the first attempt, at the best of our knowledge, which considers closed-end formulas for VolTarget-linked options. Our underlying environment can be compared to the one considered by Zagst et al. in [56], where the risky asset is assumed to evolve as in the Black-Scholes model; we extend it through considering a generalized geometric Brownian motion framework. This choice was mainly motivated by the evidence found in the papers by Binder et al. [21] and by Hilliard and Li [65], in which they heuristically showed that the stock market volatilities are mainly driven by the underlying asset price, the risk-free interest rate, the risk premium on equity and the ratio of expected returns. Since our aim is not to study the effect of variations of the interest rate and we are considering risk-neutral pricing formulas, the principal factor in determining the level of volatility of the stock price is the stock price value itself. In particular in [65] they provided a measure for the implied volatility based on the underlying asset price changes, showing that indeed, even for the implied volatility hedging approach, including the underlying market price enables to outperform traditional methods not including the relationship of the volatility with respect of the underlying price. In our study we focus on the derivation of a closed-end formula for call and put options linked to VolTarget strategies. We also consider the associated Greeks, deriving closed-end expressions for key hedging parameters of options linked to Target Volatility Strategies. Such a novel result constitutes a key point for any practitioner aiming at concretely exploiting options linked to the VolTarget portfolios.

Chapter 9 is organized as follows: in Section 9.1 we analyze VolTarget portfolios when the risky-asset dynamics is described by a generalized geometric Brownian motion. We treat the evaluation problem for options that have, as underlying, VolTarget portfolios determined by standard VolTarget strategies, preserving fixed volatility in time. In section 9.3 we consider a modification of the VolTarget strategy, which is placing an upper bound to the leverage effect caused by the VolTarget strategy. In this case we consider a risky asset evolving as a geometric Brownian motion with time-dependent drift and volatility. For both strategies exact formulas for the price of call and put options are presented. In section 9.4 we analyze the sensibility of the prices of options written on VolTarget portfolios with respect to volatility and risky asset value. We give emphasis on the analysis of the Greeks (Vega, Delta and Gamma) with respect to changes in the underlying volatility, also providing several graphs to better illustrate the obtained results. In section 9.5 we enrich our study by presenting a numerical scheme via Euler-Maruyama and Milstein discretization to simulate paths of a VolTarget portfolio, assuming that the risky asset is described by a Heston model.

Chapter 9

VolTarget strategies in generalized GBM environment

Let us start by considering a framework similar to the pioneering paper by Merton [90]. This means that throughout this chapter we are going to consider a market in which there are two investment opportunities: a *risk-free asset*, also referred as money market or Government bond or simply bond, and a *risky underlying asset*, also called stock or share. Moreover we assume that the randomness of the underlying asset is described by Black-Scholes-Merton stochastic differential equations and that there exist continuously-trading perfect markets where the agents are not subjected to any transaction costs to trade the risky asset for the risk-free asset and vice versa.

9.1 The model

Let $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$ be a filtered, complete probability space, with right-continuous filtration, supporting a Brownian motion W, and consider a market consisting of two investment opportunities: a risky asset $\{S(t)\}_{t\geq 0}$, and a risk-free asset $\{B(t)\}_{t\geq 0}$, evolving as a stochastic process satisfying the generalized geometric Brownian motion and a deterministic function:

$$\begin{aligned} \mathrm{d}S(t) &= S(t)\left(\mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t)\right), \\ \mathrm{d}B(t) &= r\,B(t)\,\mathrm{d}t\,, \end{aligned} \tag{9.1.1}$$

for each $t \geq t_0$, where $t_0 \geq 0$ is the starting time, W is a Brownian motion $\mathcal{F}(t)$-adapted, $r \in \mathbb{R}^+$ is a positive constant representing the risk-free rate, and μ and σ are adapted to $\{\mathcal{F}^W\}$, the natural filtration generated by the Brownian motion, and represent the mean rate of return and the percentage volatility of the risky asset respectively. Let $s, b \in \mathbb{R}_+$ be the values at time t_0 for the risky and risk-free assets.

Moreover, consider an investor holding a portfolio, starting with a positive position x invested in the risk-free asset and a positive position y invested in the risky asset, and assume that she is able to transfer its capitals from an investment to another without paying any transaction costs. Therefore let us introduce the processes L representing the cumulative amount of risk-free asset sold in order to buy risky asset, and M the process representing the cumulative amount of risky asset sold in order to buy the risk-free asset. Both L and M are assumed to be non-negative, non-decreasing and càdlàg.

Finally, the portfolio value can be represented continuously in time by the couple $(X(t), Y(t))_{t \geq t_0}$, starting at $X(t_0) = x$, $Y(t_0) = y$, with (X, Y) representing the amount of capital invested in the risk-free asset and in the risky asset, respectively, and evolving according to the following stochastic differential equations

$$\mathrm{d}X(t) = r\, X(t)\, \mathrm{d}t + \mathrm{d}M(t) - \mathrm{d}L(t)\,, \tag{9.1.2}$$

$$\mathrm{d}Y(t) = \mu(t)\, Y(t)\, \mathrm{d}t + \sigma(t)\, Y(t)\, \mathrm{d}W(t) + \mathrm{d}L(t) - \mathrm{d}M(t)\,. \tag{9.1.3}$$

For the moment in our work we are not introducing neither proportional or fixed transaction costs, but it is a possible extension for future studies. In such cases, continuous re-balancing would cause non-negligible expenses to the investor in the portfolio aiming to preserve a fixed volatility and it has to be considered a volatility target interval, instead of a punctual volatility target.

Denote the total portfolio value of the investor at time $t > t_0$ by $V(t)$, and let $\alpha(t)$ denote the percentage of portfolio invested at time t in the risky asset assumed to be adapted predictable càdlàg processes, while $1 - \alpha(t)$ will denote the portfolio weight invested in the risk-free-asset at time t. In other words we define $V(t) = X(t) + Y(t)$ and $\alpha(t) = \frac{Y(t)}{X(t)+Y(t)}$. Since the investments evolve according to (9.1.2) and (9.1.3), by substituting the risky asset dynamics (9.1.1), we derive the dynamics of the portfolio value process:

$$\begin{cases} \mathrm{d}V(t) = V(t)\big\{\big(\alpha(t)\,(\mu(t) - r) + r\big)\, \mathrm{d}t + \alpha(t)\, \sigma(t)\, \mathrm{d}W(t)\big\}, \quad t > t_0\,, \\ V(0) = x + y =: v, \end{cases} \tag{9.1.4}$$

where α is controlled by the investor and adapted to the filtration $\mathcal{F}$. We make two remarks: first of all notice that the portfolio whose value is determined by (9.1.4) is self-financing, i.e. the dynamics of (9.1.4) are equivalent to

$$\mathrm{d}V(t) = V(t)\left(\alpha(t)\,\frac{\mathrm{d}S(t)}{S(t)} + (1-\alpha(t))\,\frac{\mathrm{d}B(t)}{B(t)}\right), \tag{9.1.5}$$

moreover notice that V is a Markovian portfolio and that a priori we do not know the future value of the wealth process, which indeed is stochastic.

Now that the mathematical framework is settled, we can turn to the Volatility Target (VT, in short) investment strategy. The VT mechanism is a dynamic asset allocation process in which the amount invested in the risky asset is determined by a pre-defined volatility target level, denoted by $\widehat{\sigma}$, and the volatility of the underlying risky asset, $\sigma(t)$, see (9.1.1). By shifting dynamically between the two investment opportunities evolving accordingly to equations (9.1.1), the investor aims to preserve a constant volatility level of the resulting portfolio $V_{\widehat{\sigma}}(t)$, which can be used as underlying of derivatives, e.g. European call options. We resume this notion in the following definition:

Definition 9.1.1 (VolTarget portfolio) Consider the stochastic process $V_{\widehat{\sigma}}$ evolving according to

$$V_{\widehat{\sigma}}(t) = X_{1-\widehat{\alpha}}(t) + Y_{\widehat{\alpha}}(t), \quad t \geq t_0. \tag{9.1.6}$$

where $Y_{\widehat{\alpha}} = \widehat{\alpha}\, V_{\widehat{\sigma}}$ and $X_{\widehat{\alpha}} = (1-\widehat{\alpha})\, V_{\widehat{\sigma}}$, and by $\widehat{\alpha}$ we meant the proportion of portfolio value $\alpha(t)$ dynamically invested in the risky asset. We say that $V_{\widehat{\sigma}}$ is a *VolTarget portfolio* if it is self-financing and the weight process is preserving a constant volatility equal to $\widehat{\sigma}$, where X and Y represent the amount of capital invested in the risk-free and risky asset and evolve according to (9.1.2) and (9.1.3), respectively.

We want to determine explicitly the equation for the control which preserves a fixed volatility to the portfolio process (9.1.4).

Proposition 9.1.2 *For $\widehat{\alpha}(t) = \widehat{\sigma}/\sigma(t)$ we have that the process whose dynamics are given by* (9.1.6) *is a VolTarget portfolio.*

Proof. By (9.1.4), for $\alpha(t) = \widehat{\alpha}(t)$, we have

$$\mathrm{d}V_{\widehat{\sigma}}(t) = V_{\widehat{\sigma}}(t)\left(\left(\frac{\widehat{\sigma}}{\sigma(t)}\,(\mu(t)-r)+r\right)\,\mathrm{d}t + \widehat{\sigma}\,\mathrm{d}W(t)\right), \tag{9.1.7}$$

i.e. by Definition 9.1.1 we reach our thesis. □

Notice that we are not considering the equation involving the underlying (9.1.1) and the amounts of capital that have to be invested in the risky and risk-free asset respectively. In Proposition 9.1.2 we saw that in order to obtain a VT portfolio the investor has to keep this proportion inversely proportional to the actual value of the volatility rate of the risky asset, i.e. equal to $\widehat{\alpha}(t) = \widehat{\sigma}/\sigma(t)$, which is adapted to the filtration generated by the random component in the asset's stochastic differential equation.

9.2 Option pricing

Let $\mathcal{X} = \Phi(V_T)$ be a *contingent claim* with date to maturity T and with underlying portfolio V (therefore we have $\mathcal{X} \in \mathcal{F}_T$), where Φ is a contract function. The aim of this section is to determine an arbitrage-free price $\Pi(t; \mathcal{X})$ for this claim, sometimes also denoted as $\Pi(t; \Phi)$ or $\Pi(t)$.

Let us heuristically assume for the moment that there exists a function $F \in C^{1,2}([0, T] \times \mathbb{R}_+)$ such that

$$\Pi(t) = F(t, S(t)) \,,$$

then by the Black-Scholes equation we would have absence of arbitrage if F is solution to the following PDE

$$\begin{cases} \partial(t)F(t,s) + r\,s\,\partial(s)F(t,s) + \frac{1}{2}\,s^2\,\widehat{\sigma}^2\,\partial^2_{ss}F(t,s) - r\,F(t,s) & = 0\,, \\ F(T,s) & = \Phi(s)\,, \end{cases} \tag{9.2.1}$$

for $t \in [0, T]$ and $s \in \mathbb{R}_+$. In the next subsections we are going to remove the heuristic assumption and derive some pricing formulas for contingent claims written on the VolTarget portfolio. One can simply notice that the PDE (9.2.1) can be solved à la Feynman-Kač, i.e.

$$F(t,s) = e^{-r\,(T-t)}\,\mathbb{E}^{\mathbb{Q}}_{t,s}[N(S_T)], \quad \text{for } t \in [0,T], s > 0,$$

where we are considering the expectation with respect to the unique risk-neutral measure $\mathbb{Q}$ conditioned by $S(t) = s$, see e.g. [18, Ch. 14] for what concerns the existence of a unique risk-neutral measure.

9.2.1 Risk Neutral Valuation

We consider now the unique equivalent martingale measure $\mathbb{Q}$, i.e. the unique measure under which $S(t)/B(t)$ is a local martingale. Let $W^{\mathbb{Q}}$ be a Brownian motion under

$\mathbb{Q}$, then by Girsanov theorem we have that the risky asset process S satisfies the following SDE

$$\mathrm{d}S(t) = S(t)\left(r\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W^{\mathbb{Q}}(t)\right). \tag{9.2.2}$$

The underlying risky asset is governed by a geometric Brownian motion with dynamics given by equation (9.2.2), therefore applying Itô formula to $\log(S(t))$, we obtain

$$S(t) = S(0)\,\exp\left\{\int_0^t (r - \sigma(s)^2/2)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}W^{\mathbb{Q}}(s)\right\}.$$

Since the volatility in equation (9.2.2) is not a function only of time, but it is adapted to W, we cannot say much about the distribution of $\log(S(t)/S(0))$, which would have been Gaussian, in the special case of deterministic volatility.

We consider now a European call option and a European put option with payoff function

$$\Phi_{\text{call}}(V_T) = (V_T - K)_+\,, \tag{9.2.3}$$

$$\Phi_{\text{put}}(V_T) = (K - V_T)_+\,, \tag{9.2.4}$$

with $T \geq t_0$ expiration time and K strike price. Through the next propositions and corollaries we will determine the price at the starting time $t_0 \geq 0$ of these kinds of options.

Proposition 9.2.1 *The price at time t_0 of a call option with payoff (9.2.3), denoted as Φ_{call}, linked to the VolTarget portfolio $V_{\widehat{\sigma}}(t)$, see equation (9.1.7), is given by the following explicit formula*

$$\Pi(t_0, \Phi_{call}(V_{\widehat{\sigma}}(T))) = v\,N\left(d_1(t_0)\right) - K\,e^{-r\,(T-t_0)}\,N\left(d_2(t_0)\right). \tag{9.2.5}$$

We recall that the proportion $\widehat{\alpha}(t)$ of portfolio value invested in the risky-asset is as the one defined in Proposition 9.1.2, N is the cumulative distribution function for the standard normal distribution, $v = V_{\widehat{\sigma}}(t_0)$ is the starting value of the portfolio and we defined the following parameters

$$d_1(t_0) = \frac{-z_{\widehat{\sigma}}(t_0) + \widehat{\sigma}\,(T - t_0)}{\sqrt{T - t_0}}\,,$$

$$d_2(t_0) = -\frac{z_{\widehat{\sigma}}(t_0)}{\sqrt{T - t_0}}\,,$$

$$z_{\widehat{\sigma}}(t_0) = \frac{1}{\widehat{\sigma}}\log\left(\frac{K}{v}\right) + \left(\frac{\widehat{\sigma}}{2} - \frac{r}{\widehat{\sigma}}\right)(T - t_0)\,.$$

Proof. Notice that, while the underlying risky asset has non constant volatility, see eq. (9.1.7), the dynamics for the VT portfolio is simpler, and indeed we have the following explicit formula

$$V_{\widehat{\sigma}}(t) = v\,\exp\left(\left(r - \frac{\widehat{\sigma}^2}{2}\right)(t - t_0) + \widehat{\sigma}\,W^{\mathbb{Q}}(t - t_0)\right), \quad \text{for } t \geq t_0\,.$$

Therefore we have that $V_{\widehat{\sigma}}(T) > K$ if and only if

$$W^{\mathbb{Q}}(T - t_0) > \frac{1}{\widehat{\sigma}}\log\left(\frac{K}{v}\right) + \left(\frac{\widehat{\sigma}}{2} - \frac{r}{\widehat{\sigma}}\right)(T - t_0) =: z_{\widehat{\sigma}}(t_0)\,.$$

We denote by $f_{N(0,t)}(x)$ the probability density function of the a normal random variable with mean 0 and variance t

$$f_{N(0,t)}(x) = \frac{1}{\sqrt{2\pi\,t}}\,e^{-\frac{x^2}{2t}}\,,$$

and by $N(x)$ the cumulative distribution function of a standard normal variable. Then we have that the price of the call option on the portfolio value at time t_0 is equal to

$$\begin{aligned}
\Pi(t_0, \Phi_{\text{call}}(V_{\widehat{\sigma}}(T))) &= \mathbb{E}\left[e^{-r\,(T-t_0)}(V_{\widehat{\sigma}}(T) - K)_+ \,\middle|\, \mathcal{F}_{t_0}\right] \\
&= e^{-r(T-t_0)}\int_{z_{\widehat{\sigma}}(t_0)}^{+\infty}\left\{v\,\exp\left(\left(r - \frac{\widehat{\sigma}^2}{2}\right)(T - t_0) + \widehat{\sigma}x\right) - K\right\} f_{N(0,T-t_0)}(x)\mathrm{d}x \\
&= e^{-r\,(T-t_0)}\,v\,e^{\left(r-\frac{\widehat{\sigma}^2}{2}\right)(T-t_0)+\widehat{\sigma}^2/2\,(T-t_0)}\left(1 - N\left(\frac{z_{\widehat{\sigma}}(t_0) - \widehat{\sigma}\,(T - t_0)}{\sqrt{T - t_0}}\right)\right) \\
&\quad - K\,e^{-r\,(T-t_0)}\left(1 - N\left(\frac{z_{\widehat{\sigma}}(t_0)}{\sqrt{T - t_0}}\right)\right) \\
&= v\,N\left(\frac{-z_{\widehat{\sigma}}(t_0) + \widehat{\sigma}\,(T - t_0)}{\sqrt{T - t_0}}\right) - K\,e^{-r\,(T-t_0)}\,N\left(-\frac{z_{\widehat{\sigma}}(t_0)}{\sqrt{T - t_0}}\right).
\end{aligned}$$

□

Corollary 9.2.2 *Assuming that the risky asset dynamics follows a generalized geometric Brownian motion with $\mathcal{F}^W$-adapted drift and volatility, see equation (9.1.1), the price at time t_0 of a put option with payoff (9.2.4), denoted as Φ_{put}, linked to the VolTarget portfolio $V_{\widehat{\sigma}}(t)$, see equation (9.1.7), is given by the following explicit* *formula*

$$\Pi(t_0, \Phi_{put}(V_{\widehat{\sigma}}(T))) = K\, e^{-r\,(T-t_0)}\, N\left(-d_2(t_0)\right) - v\, N\left(-d_1(t_0)\right), \tag{9.2.6}$$

where the parameters d_1, d_2 and $z_{\widehat{\sigma}}$ are defined as in Proposition 9.2.1.

Proof. By the put-call Parity, see e.g. [110, 4.5.6], we have that the difference between the price of a call option and the price of put option with same strike price, time to expiration and underlying, is equal to difference between the actual price of the underlying (in our case the VT portfolio) and the discounted strike price, i.e.

$$\Pi(t_0, \Phi_{\text{call}}(V_{\widehat{\sigma}}(T))) - \Pi(t_0, \Phi_{\text{put}}(V_{\widehat{\sigma}}(T))) = v - K\, e^{-r\,(T-t_0)}.$$

Therefore, we obtain (9.2.6), since $N(-x) = 1 - N(x)$ for each $x \in \mathbb{R}$. □

9.3 VolTarget Strategy with maximum allowed Leverage Factor

Now we consider a more interesting strategy from a practical point of view. We introduce a parameter $L \geq 1$ determining the maximum allowed leverage of the portfolio, i.e. we force the weight process to be less or equal than the parameter L:

$$\widetilde{\alpha}(t) := \min\{L; \widehat{\sigma}/\sigma(t)\}\,. \tag{9.3.1}$$

From now on, we will distinguish the notations for standard VolTarget strategies by the one for VolTarget strategies with maximum allowed Leverage Factor, by marking the volatility and weight symbols with a hat and a tilde respectively, i.e. $\widehat{\sigma}$ and $\widehat{\alpha}$ are referred to standard VolTarget portfolios and $\widetilde{\sigma}$ and $\widetilde{\alpha}$ are referred to "Maximum Leverage" VolTarget portfolios.

This limitation is imposed in order to prohibit VolTarget strategies to finance by loans a large portion of the risky investment. The typical setup occurring in practice is $L = 2$, see the paper by Wallbaum et al. [7] for further details.

Let us compute the value of an option on the VolTarget portfolio with limited leverage and time dependent volatility, i.e. for this VT strategy we consider a particular case of equation (9.1.1):

$$\mathrm{d}S(t) = S(t)\big(\mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t)\big), \tag{9.3.2}$$

where $\mu, \sigma : \mathbb{R}^+ \to \mathbb{R}^+$ are deterministic functions of time. Here the percentage drift term could be assumed to be $\mathcal{F}^W$-adapted.

Proposition 9.3.1 *Assume that the risky asset dynamics follows a geometric Brownian motion with time-dependent drift and volatility, see equation (9.3.2). The price at time t_0 of a call option with payoff (9.2.3), denoted as Φ_{call}, linked to the VolTarget leverage portfolio $V_{\widetilde{\sigma}}(t)$, is given by the following explicit formula*

$$\Pi\left(t_0, \Phi_{call}(V_{\widetilde{\sigma}}(T))\right) = v\, N\left(\widetilde{d}_1(t_0)\right) - K\, e^{-r\,(T-t_0)}\, N\left(\widetilde{d}_2(t_0)\right), \tag{9.3.3}$$

where the proportion of portfolio value invested in the risky-asset is $\widetilde{\alpha}(t) := \min\{L; \widehat{\sigma}/\sigma(t)\}$, N is the cumulative distribution function for the standard normal distribution, $v = V_{\widetilde{\sigma}}(t_0)$ is the starting value of the portfolio and we defined the following parameters

$$\begin{aligned}
\widetilde{d}_1(t_0) &= \frac{-\widetilde{z}_{\widetilde{\sigma}}(t_0) + \varsigma(t_0)}{\sqrt{\varsigma(t_0)}}\,,\\
\widetilde{d}_2(t_0) &= -\frac{\widetilde{z}_{\widetilde{\sigma}}(t_0)}{\sqrt{\varsigma(t_0)}}\,,\\
z_{\widetilde{\sigma}}(t_0) &= \log\left(\frac{K}{v}\right) - r\,(T-t_0) + \frac{\varsigma(t_0)}{2}\,,\\
\varsigma(t_0) &= \int_{t_0}^{T} \widetilde{\sigma}(s)^2\,\mathrm{d}s\,,\\
\widetilde{\sigma}(t) &= \min\{L\,\sigma(t), \widehat{\sigma}\}\,.
\end{aligned}$$

Proof. For this strategy we have that the portfolio value has not a constant volatility and it has the following expression

$$V_{\widetilde{\sigma}}(t_0, t) = v\,\exp\left(r\,(t-t_0) - \varsigma(t_0)/2 + \int_{t_0}^{t} \min(L\,\sigma(s), \widehat{\sigma})\,\mathrm{d}W^{\mathbb{Q}}(s)\right),$$

and we have $\widetilde{W}(t-t_0) := \int_{t_0}^{t} \min(L\,\sigma(s), \widehat{\sigma})\,\mathrm{d}W^{\mathbb{Q}}(s) \sim N(0, \varsigma(t_0))$, which means that its probability density function is

$$f_{N(0,\varsigma(t_0))}(x) = \frac{1}{\sqrt{2\,\pi\,\varsigma(t_0)}}\exp\left(-\frac{x^2}{2\,\varsigma(t_0)}\right).$$

Therefore we have that $V_{\widetilde{\sigma}}(T) > K$ iff

$$\widetilde{W}(T-t_0) > \log\left(\frac{K}{v}\right) - r\,(T-t_0) + \frac{\varsigma(t_0)}{2} =: z_{\widetilde{\sigma}}(t_0)\,,$$

and have we have that the considered option value equals to

$$\begin{aligned}
\Pi\left(t_0, \Phi_{\text{call}}(V_{\widetilde{\sigma}}(T))\right) &= \mathbb{E}\left[e^{-r\,(T-t_0)}(V_{\widetilde{\sigma}}(T)-K)_+ \middle| \mathcal{F}_{t_0}\right] \\
&= e^{-r(T-t_0)}\int_{z_{\widetilde{\sigma}}}^{+\infty}\Big\{v\exp\Big(r(T-t_0)-\varsigma(t_0)/2+x\Big)-K\Big\}f_{N(0,\varsigma(t_0))}(x)\mathrm{d}x \\
&= e^{-r\,(T-t_0)}\,v\,e^{r\,(T-t_0)-\varsigma(t_0)/2+\varsigma(t_0)/2}\left(1-N\left(\frac{z_{\widetilde{\sigma}}(t_0)-\varsigma(t_0)}{\sqrt{\varsigma(t_0)}}\right)\right) \\
&\qquad - K\,e^{-r\,(T-t_0)}\left(1-N\left(\frac{z_{\widetilde{\sigma}}(t_0)}{\sqrt{\varsigma(t_0)}}\right)\right) \\
&= v\,N\left(\frac{-z_{\widetilde{\sigma}}(t_0)+\varsigma(t_0)}{\sqrt{\varsigma(t_0)}}\right) - K\,e^{-r\,(T-t_0)}\,N\left(-\frac{z_{\widetilde{\sigma}}(t_0)}{\sqrt{\varsigma(t_0)}}\right).
\end{aligned}$$

□

Remark 9.3.2 Notice that the price of this call option depends on the future volatility, but, since it is deterministic, it is not an issue, and indeed we have obtained exact formulas.

Corollary 9.3.3 *Assume that the risky asset dynamics follows a geometric Brownian motion with time-dependent drift and volatility, see equation (9.3.2). Then the price at time t_0 of a put option with payoff (9.2.4), denoted as Φ_{put}, linked to the VolTarget leverage portfolio $V_{\widetilde{\sigma}}(t)$, is given by the following explicit formula*

$$\Pi\left(t_0, \Phi_{put}(V_{\widetilde{\sigma}}(T))\right) = K\,e^{-r\,(T-t_0)}\,N\left(-\widetilde{d}_2(t_0)\right) - v\,N\left(-\widetilde{d}_1(t_0)\right) \tag{9.3.4}$$

where $\widetilde{d}_1$ and $\widetilde{d}_2$ are defined as in Proposition 9.3.1, and the proportion of portfolio value invested in the risky asset is $\widetilde{\alpha}(t) := \min\{L; \widehat{\sigma}/\sigma(t)\}$.

Proof. Direct consequence of the put-call Parity. □

9.4 Greeks

In this section we will move on the quantitative study of the prices of options on a VolTarget portfolio in continuous time. In other words, we will explicitly derive the *Greeks*, which are known in mathematical finance as the quantities representing the sensitivity of prices of derivatives, such as options, to a change in underlying parameters on which the value of the portfolio of financial instruments is dependent.

In what follows we are going to consider a risky asset evolving as in the Black-Scholes model. Therefore, the price formulas (9.2.5), (9.2.6), (9.3.3) and (9.3.4) for call and put options, with VT underlying portfolios, can be reduced to

$$
\begin{aligned}
\Pi\left(t_0, \Phi_{\text{call}}(V_{\widehat{\sigma}}(T))\right) =& v\, N(d_1) - K\, e^{-r\,(T-t_0)}\, N(d_2)\,,\\
\Pi\left(t_0, \Phi_{\text{call}}(V_{\widetilde{\sigma}}(T))\right) =& \begin{cases} v\, N(\widetilde{d}_1) - K\, e^{-r\,(T-t_0)}\, N(\widetilde{d}_2) & \text{for } \sigma < \widehat{\sigma}/L\,,\\ v\, N(d_1) - K\, e^{-r\,(T-t_0)}\, N(d_2) & \text{for } \sigma > \widehat{\sigma}/L\,,\end{cases}\\
\Pi\left(t_0, \Phi_{\text{put}}(V_{\widehat{\sigma}}(T))\right) =& K\, e^{-r\,(T-t_0)}\, N(-d_2) - v\, N(-d_1)\,,\\
\Pi\left(t_0, \Phi_{\text{put}}(V_{\widetilde{\sigma}}(T))\right) =& \begin{cases} K\, e^{-r\,(T-t_0)}\, N(-\widetilde{d}_2) - v\, N(-\widetilde{d}_1) & \text{for } \sigma < \widehat{\sigma}/L\,,\\ K\, e^{-r\,(T-t_0)}\, N(-d_2) - v\, N(-d_1) & \text{for } \sigma > \widehat{\sigma}/L\,,\end{cases}
\end{aligned}
$$

where

$$
\begin{aligned}
d_1 =& \frac{1}{\widehat{\sigma}\,\sqrt{T-t_0}}\,\left(\log(v/K) + (r + \widehat{\sigma}^2/2)\,(T-t_0)\right)\,,\\
d_2 =& \frac{1}{\widehat{\sigma}\,\sqrt{T-t_0}}\,\left(\log(v/K) + (r - \widehat{\sigma}^2/2)\,(T-t_0)\right)\,,\\
\widetilde{d}_1 =& \frac{1}{\sigma\, L\,\sqrt{T-t_0}}\,\left(\log(v/K) + (r + L^2\,\sigma^2/2)\,(T-t_0)\right)\,,\\
\widetilde{d}_2 =& \frac{1}{\sigma\, L\,\sqrt{T-t_0}}\,\left(\log(v/K) + (r - L^2\,\sigma^2/2)\,(T-t_0)\right)\,.
\end{aligned}
$$

9.4.1 Vega

Since VolTarget portfolios are meant to preserve a fixed volatility the most representative Greek value is the *Vega* value, i.e. the sensitivity of the option price with respect to the volatility of the risky asset.

Proposition 9.4.1 *The Vega of a call and put option, with payoff* (9.2.3) *and* (9.2.4)*, on the VolTarget portfolio with weight strategies* $\widehat{\alpha} = \widehat{\sigma}/\sigma$ *and* $\widetilde{\alpha} := \min\{L; \widehat{\sigma}/\sigma\}$ *are respectively given by*

$$
\nu_{\{\Phi_{call}, V_{\widehat{\sigma}}\}} = \partial_\sigma \Pi\left(t_0, \Phi_{call}(V_{\widehat{\sigma}}(T))\right) = 0\,,
$$

$$
\nu_{\{\Phi_{call}, V_{\widetilde{\sigma}}\}} = \partial_\sigma \Pi\left(t_0, \Phi_{call}(V_{\widetilde{\sigma}}(T))\right) = \begin{cases} \frac{v}{\sqrt{2\pi}}\, \exp\left(-\frac{\widetilde{d}_1^2}{2}\right) L\,\sqrt{T-t_0} & \text{for } \sigma < \frac{\widehat{\sigma}}{L},\\ 0 & \text{for } \sigma > \frac{\widehat{\sigma}}{L},\end{cases} \tag{9.4.1}
$$

with $\nu_{\{\Phi_{put},V_{\widehat{\sigma}}\}} = \nu_{\{\Phi_{call},V_{\widehat{\sigma}}\}}$ *and* $\nu_{\{\Phi_{put},V_{\widetilde{\sigma}}\}} = \nu_{\{\Phi_{call},V_{\widetilde{\sigma}}\}}$*, where*

$$\widetilde{d}_1 = \frac{\log(v/K) + \left(r + \frac{L^2\sigma^2}{2}\right)(T - t_0)}{L\,\sigma\,\sqrt{T - t_0}}\,.$$

Proof. Let us consider the VT leverage strategy for the case in which $\sigma < \frac{\widehat{\sigma}}{L}$, then we recall that the price of the call option can simplified as

$$\Pi\,(t_0, \Phi_{\text{call}}(V_{\widetilde{\sigma}}(T))) = v\,N(\widetilde{d}_1) - K\,\exp(-r\,(T-t))\,N(\widetilde{d}_2)\,,$$

where

$$\begin{aligned}\widetilde{d}_1 &= \frac{-\log(K/v) + (r + L^2\,\sigma^2/2)\,(T-t)}{L\,\sigma\,\sqrt{T-t_0}}\,,\\ \widetilde{d}_2 &= \frac{-\log(K/v) + (r - L^2\,\sigma^2/2)\,(T-t)}{L\,\sigma\,\sqrt{T-t_0}}\,.\end{aligned}$$

Then computing the partial derivative with respect to σ, we obtain

$$\begin{aligned}\partial_\sigma\Pi\,(t_0, \Phi_{\text{call}}(V_{\widetilde{\sigma}}(T))) &= \frac{1}{\sqrt{2\pi}}\left(v\,e^{-\widetilde{d}_1^2/2}\left(L\,\sqrt{T-t_0} - \frac{\widetilde{d}_1}{\sigma}\right)\right.\\ &\qquad \left. + K\,e^{-\widetilde{d}_2^2/2 - r\,(T-t_0)}\left(L\,\sqrt{T-t_0} + \frac{\widetilde{d}_2}{\sigma}\right)\right)\\ &= \frac{1}{\sqrt{2\pi}}\,v\,e^{-\widetilde{d}_1^2/2}\left(2\,L\,\sqrt{T-t_0} - \frac{\widetilde{d}_1 - \widetilde{d}_2}{\sigma}\right)\\ &= \frac{v}{\sqrt{2\pi}}\,\exp\left(-\frac{(\widetilde{d}_1)^2}{2}\right)L\,\sqrt{T-t_0}\,,\end{aligned}$$

where the second row steams from the identity $\exp(\frac{\widetilde{d}_2^2 - \widetilde{d}_1^2}{2}) = \frac{v}{K}\exp(r\,(T-t_0))$, and the last by $\widetilde{d}_1 - \widetilde{d}_2 = L\,\sigma\,\sqrt{T-t_0}$.

Similar calculations work also for the Vega of put options. □

Remark 9.4.2 Notice that, for $\sigma < \widehat{\sigma}/L$, we have that VT leverage call and the standard leverage call share the same price, which can be expressed in terms of the standard call option price, denoted as $\Pi_S := \Pi\,(t_0, \Phi_{\text{call}}(S(T)))$; i.e. stating the dependence with respect the risky asset's volatility

$$\Pi_{V_{\widetilde{\sigma}}}(\sigma) = \Pi_S(L\,\sigma)\,,$$

and therefore, computing the partial derivative with respect to the volatility, we obtain

$$\nu_{\{\Phi_{\text{call}}, V_{\widehat{\sigma}}\}} = \partial_\sigma \Pi_{V_{\widehat{\sigma}}}(\sigma) = L\, \partial_\sigma \Pi_S(L\,\sigma)\,,$$

which is exactly as expressed in equation (9.4.1).

In Figure 9.1 we compared the graphs of a Vega for a call option written on a portfolio adopting a VT leverage strategy and the Vega for a standard call option. In the left graph in Figure 9.1, the Vegas taken into account are for at-the-money options, and one can notice that, while for volatilities higher than $\sigma > \frac{\widehat{\sigma}}{L}$, the VT leverage Vega is null, for "small" volatilities the VT leverage Vega is even higher than the Vega for standard call options. In the right graph in Figure 9.1 the comparison takes into account the sensitivity of Vega with respect to the portfolio value. Notice that, while for a standard call option the highest Vega is reached for the underlying share's value equal

$$v^* = s^* = K\, e^{-(T-t_0)\,(r-\frac{\sigma^2}{2})},$$

for the VT leverage call option, it is reached in

$$v^* = K\, e^{-(T-t_0)\,(r-\frac{L^2\sigma^2}{2})}.$$

Finally, in Figure 9.2, we summarized the dependence of Vega with respect to both the volatility and the underlying portfolio value.

9.4.2 Delta

Before dealing with the Delta of options written on VT portfolios, it is worth starting analyzing the sensitivity of the VT portfolio with respect to small changes in the risky asset price. To perform this, we write the VT portfolio dynamics as

$$\mathrm{d}V_{\widehat{\sigma}}(t) = \varphi_{\widehat{\sigma}}(t)\,\mathrm{d}S(t) + \psi_{\widehat{\sigma}}(t)\,\mathrm{d}B(t)\,, \tag{9.4.2}$$

where we defined φ_S and φ_B as the instantaneous number of shares and bonds held in the portfolio. By the self-financing equation (9.1.5), we have that

$$\varphi_{\widehat{\sigma}}(t) = \frac{V(t)\,\widehat{\alpha}(t)}{S(t)}\,,$$
$$\psi_{\widehat{\sigma}}(t) = \frac{V(t)\,(1-\widehat{\alpha}(t))}{B(t)}\,,$$

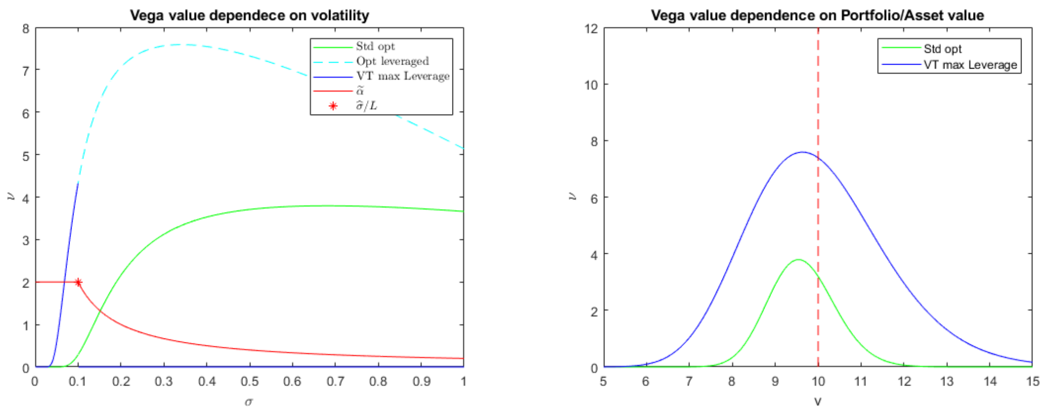

Figure 9.1: The plots represent the behavior of Vega values of an option written on the VolTarget portfolio, with maximum allowed leverage given by the weight strategy $\widetilde{\alpha} = \min(L, \widehat{\sigma}/\sigma)$, highlighting its dependence on the values of the volatility σ (left) and on the values of v (right). The parameters are set up as $r = 5\%, v = 12, K = 10, t_0 = 0, T = 1, \widehat{\sigma} = 20\%, L = 2$. For the volatility dependence (left), the "VolTarget Maximum Leverage" Vega line (in blue) is also compared with the dotted line of a hypothetical portfolio holding L times its wealth in the risky asset (in cyan) and the Vega for a standard call option whose underlying is simply the risky asset. Instead, for the portfolio initial value dependence (right), we considered $\sigma = 0.08$, i.e. $\sigma < \widehat{\sigma}/L$. The "VolTarget Maximum Leverage" Vega line (in blue) is the same as the Vega of a hypothetical portfolio holding L times its wealth in the risky asset. This line is compared with the Vega of a portfolio investing all its capital in the risky asset. We remark the fact that here we obtained that the Vega value for the VT leverage option was greater than the one for the standard option, since we were considering a relatively small volatility (less than $\widehat{\sigma}/L$); instead, if we would have considered a volatility greater than $\widehat{\sigma}/L$, the Vega value for the option written on the VolTarget Maximum Leverage would have been identically zero.

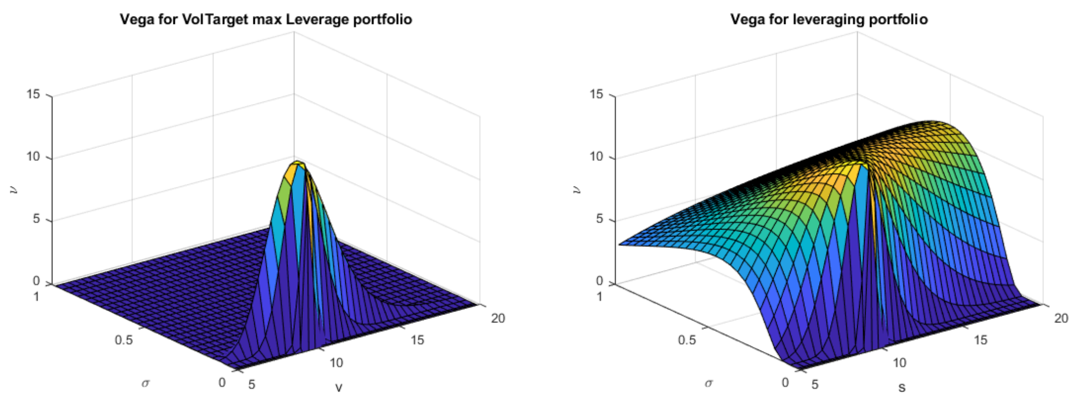

Figure 9.2: The two surfaces are the Vega for a portfolio adopting a VolTarget maximum leverage strategy (left figure) and the Vega for a portfolio investing L times its wealth v in the risky asset (right figure). One can notice that the VT leverage strategy hedges well the portfolio against volatility variations when the volatility is high (higher than $\widetilde{\sigma}/L$). The parameters are set as $r = 5\%, t_0 = 0, T = 1, \widehat{\sigma} = 20\%, L = 2, K = 10$.

which means that the Delta of the VT portfolio is

$$\Delta_{V_{\widehat{\sigma}}} = \frac{V(t)\,\widehat{\sigma}}{S(t)\,\sigma}\,. \tag{9.4.3}$$

Remark 9.4.3 Let us digress on some implications resulting by Itô calculus. Notice that the price of an option written on the VT portfolio may be equivalently determined solely by the dynamics of the VT portfolio and the actual time, or by the risky asset dynamics, the bond dynamics and the actual time; i.e. we may denote the price of a generic option on a VT portfolio as $\Pi(t, V)$ or $\Pi(t, S, B)$. Therefore, exploiting the first notation, by equation (9.4.2) and Itô-Doeblin formula, see [110, Ch. 4], we have

$$\begin{aligned} \mathrm{d}\Pi(t,V) =& \partial_t \Pi(t,V)\,\mathrm{d}t + \partial_V \Pi(t,V)\,\mathrm{d}V_t + \frac{1}{2}\partial^2_{VV}\Pi(t,V)\,\mathrm{d}[V,V]_t \\ =& \partial_t \Pi(t,V)\,\mathrm{d}t + \partial_V \Pi(t,V)\,\big(\varphi(t)\,\mathrm{d}S_t + \Psi(t)\,\mathrm{d}B_t\big) \\ & + \frac{1}{2}\partial^2_{VV}\Pi(t,V)\,\varphi(t)^2\,\mathrm{d}[S,S]_t\,, \end{aligned} \tag{9.4.4}$$

where by $[V, V]$ we denote, as usual, the quadratic variation of the stochastic process V, see e.g. [110, 3.4.2].

Instead, for the expression of the option price as function of time, risky asset price and bond price, by Itô-Doeblin formula, we obtain

$$\mathrm{d}\Pi(t,S,B) = \partial_t\Pi(t,S,B)\,\mathrm{d}t + \partial_S\Pi(t,S,B)\,\mathrm{d}S_t + \partial_B\Pi(t,S,B)\,\mathrm{d}B_t + \frac{1}{2}\partial^2_{SS}\Pi(t,S,B)\,\mathrm{d}[S,S]_t, \quad (9.4.5)$$

therefore, combining equations (9.4.4) and (9.4.5), we derive a simpler expression for Delta and Gamma of an option on VT portfolios

$$\partial_S\Pi(t,S,B) = \partial_V\Pi(t,V)\,\varphi(t)\,,$$
$$\partial^2_{SS}\Pi(t,S,B) = \partial^2_{VV}\Pi(t,V)\,\varphi(t)^2\,.$$

Proposition 9.4.4 *The Delta of a European call option with payoff* (9.2.3) *on the VolTarget portfolio with weight strategies* $\widehat{\alpha} = \widehat{\sigma}/\sigma$ *and* $\widetilde{\alpha} := \min\{L;\widehat{\sigma}/\sigma\}$ *are respectively given by*

$$\Delta_{\{\Phi_{call},V_{\widehat{\sigma}}\}} = \partial_S\Pi\left(t_0,\Phi_{call}(V_{\widehat{\sigma}}(T))\right) = \frac{v\,\widehat{\sigma}}{s\,\sigma}\,N(d_1)\,, \quad (9.4.6)$$

$$\Delta_{\{\Phi_{call},V_{\widetilde{\sigma}}\}} = \partial_S\Pi\left(t_0,\Phi_{call}(V_{\widetilde{\sigma}}(T))\right) = \begin{cases} \frac{L\,v}{s}\,N(\widehat{d}_1) & \textit{for } \sigma < \frac{\widehat{\sigma}}{L}\,, \\ \frac{v\,\widehat{\sigma}}{s\,\sigma}\,N(d_1) & \textit{for } \sigma > \frac{\widehat{\sigma}}{L}\,, \end{cases} \quad (9.4.7)$$

where

$$d_1 = \frac{\log(v/K) + \left(r + \frac{\widehat{\sigma}^2}{2}\right)(T-t_0)}{\widehat{\sigma}\,\sqrt{T-t_0}}\,,$$
$$\widetilde{d}_1 = \frac{\log(v/K) + \left(r + \frac{L^2\,\sigma^2}{2}\right)(T-t_0)}{L\,\sigma\,\sqrt{T-t_0}}\,.$$

While the Delta of a European put option with payoff (9.2.4) *is*

$$\Delta_{\{\Phi_{put},V_{\widehat{\sigma}}\}} = \partial_S\Pi\left(t_0,\Phi_{put}(V_{\widehat{\sigma}}(T))\right) = \frac{v\,\widehat{\sigma}}{s\,\sigma}\,(N(d_1)-1)\,, \quad (9.4.8)$$

$$\Delta_{\{\Phi_{put},V_{\widetilde{\sigma}}\}} = \partial_S\Pi\left(t_0,\Phi_{put}(V_{\widetilde{\sigma}}(T))\right) = \begin{cases} \frac{L\,v}{s}\,(N(\widehat{d}_1)-1) & \textit{for } \sigma < \frac{\widehat{\sigma}}{L}\,, \\ \frac{v\,\widehat{\sigma}}{s\,\sigma}\,(N(d_1)-1) & \textit{for } \sigma > \frac{\widehat{\sigma}}{L}\,. \end{cases} \quad (9.4.9)$$

Proof. By Itô calculus' chain rule (see Remark 9.4.3)

$$\Delta_{\{\Phi_{\mathrm{call}},V_{\widehat{\sigma}}\}} = \partial_V\Pi\left(t_0,\Phi_{\mathrm{call}}(V_{\widehat{\sigma}}(T))\right)\,\Delta_{V_{\widehat{\sigma}}}\,. \quad (9.4.10)$$

Moreover, by simple calculation, we have

$$\begin{aligned}\partial_V \Pi\,(t_0, \Phi_{\text{call}}(V_{\widehat{\sigma}}(T))) &= N(d_1) + v\,N'(d_1)\,\partial_v d_1 - K\,e^{-r,(T-t_0)} N'(d_2)\,\partial_v d_2 \\ &= N(d_1) + \frac{1}{\sqrt{2\pi}\,\sqrt{T-t_0}\,\widehat{\sigma}^2\,v^2}\left(v\,e^{-\frac{1}{2}d_1^2} - K\,e^{-r\,(T-t_0)} e^{-\frac{1}{2}d_2^2}\right) \\ &= N(d_1) + \frac{e^{-\frac{1}{2}d_1^2}}{\sqrt{2\pi}\,\sqrt{T-t_0}\,\widehat{\sigma}^2\,v}\left(1 - \frac{K}{v}\,e^{-r\,(T-t_0)} e^{-\frac{1}{2}(d_2^2-d_1^2)}\right) = N(d_1)\,,\end{aligned} \tag{9.4.11}$$

where the last equality is reached since $e^{-\frac{1}{2}(d_2^2-d_1^2)} = v/K\,e^{r\,(T-t_0)}$. Therefore, substituting (9.4.3) and (9.4.11) in equation (9.4.10), we obtain (9.4.6).

The same arguments works for $\Delta_{V_{\widetilde{\sigma}}}$, with the difference that in this case we have

$$\Delta_{V_{\widetilde{\sigma}}} = \frac{L\,v}{s}\,, \quad \text{for } \sigma < \frac{\widehat{\sigma}}{L}\,,$$

since for $\sigma < \frac{\widehat{\sigma}}{L}$, $\widetilde{\alpha} = L$.

For the put option case, we have that $\partial_V \Pi\,(t_0, \Phi_{\text{call}}(V_{\widehat{\sigma}}(T))) = N(d_1) - 1$. Therefore, by similar calculations, we obtain equation (9.4.8). □

In Figure 9.3 we compared the Deltas for calls and puts written on a VolTarget portfolio with calls and puts written on the risky asset. One can notice that Deltas for VT options present an asymptotic behavior for both low and high volatilities. This is because, for extreme low volatilites, the VT portfolio finances a great amount of shares through short selling the risk-free asset, while, for extreme high volatilities, the VT portfolio invests only a small proportion of its value in the risky asset. See Remark 9.4.3 for the mathematical explanation of this effect.

For VT leverage portfolios the analysis of the Deltas with respect to the volatility of the risky asset is not significantly different than the Deltas for standard options, since it suffices to consider the change of variable $\sigma_L = L\,\sigma$ in the standard Delta value and to boost the latter by the maximum leverage parameter L.

9.4.3 Gamma

The computation of the Gamma for options on VT portfolios follows similar proceedings to the ideas beyond the computation of the Delta performed in Subsection 9.4.2, and in particular Remark 9.4.3.

Proposition 9.4.5 *The Gamma of an option with payoff* (9.2.3) *on the VolTarget portfolio with weight strategies* $\widehat{\alpha} = \widehat{\sigma}/\sigma$ *and* $\widetilde{\alpha} := \min\{L; \widehat{\sigma}/\sigma\}$ *are respectively given*

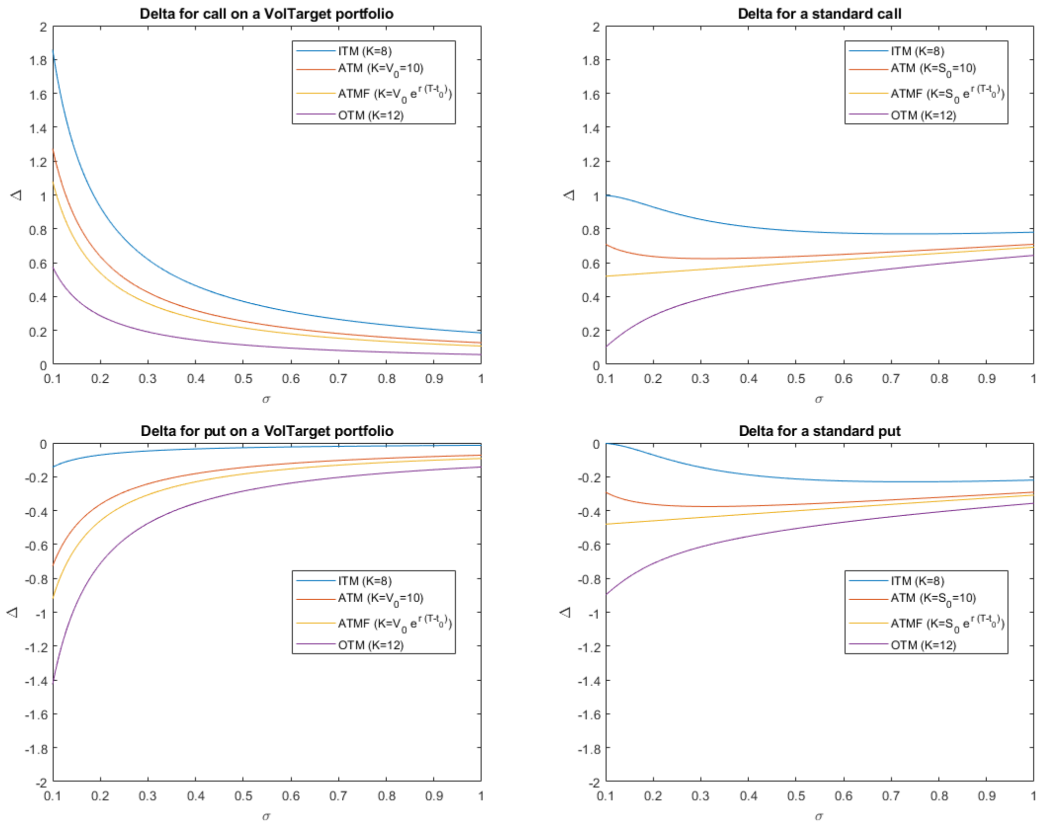

Figure 9.3: The right graphs represent the behavior of the Delta of standard call (top figure) and put (bottom figure) options with respect to different volatility values. The left graphs represent the Delta for call and put options written on VolTaget portfolios. We considered strike prices K in order to obtain an "in-the-money" option, an "at-the-money" option, an "at-the-money-forward" option and an "out-of-the-money" option. Notice that the Delta for the VT options exhibits two asymptotes: one vertical, corresponding to null volatility, and one horizontal, corresponding to a volatility value tending to infinity. The parameters are fixed as $s = v = 10$, $\widehat{\sigma} = 0.2$, $\mu = 8\%$, $r = 5\%$, $T = 1$, $t_0 = 0$ and the volatilities start at $\sigma = 0.1$.

by

$$\Gamma_{\{\Phi_{call}, V_{\widehat{\sigma}}\}} = \partial^2_{SS}\Pi\left(t_0, \Phi_{call}(V_{\widehat{\sigma}}(T))\right) = \frac{v\,\widehat{\sigma}}{s^2\,\sigma^2\,\sqrt{T-t_0}}\, f_{N(0,1)}(d_1)\,, \tag{9.4.12}$$

$$\Gamma_{\{\Phi_{call}, V_{\widetilde{\sigma}}\}} = \partial^2_{SS}\Pi\left(t_0, \Phi_{call}(V_{\widetilde{\sigma}}(T))\right) = \begin{cases} \frac{L\,v}{s^2\,\sigma\,\sqrt{T-t_0}}\, f_{N(0,1)}(\widehat{d}_1) & \textit{for } \sigma < \frac{\widehat{\sigma}}{L}\,, \\ \frac{v\,\widehat{\sigma}}{s^2\,\sigma^2\,\sqrt{T-t_0}}\, f_{N(0,1)}(d_1) & \textit{for } \sigma > \frac{\widehat{\sigma}}{L}\,, \end{cases} \tag{9.4.13}$$

with $\Gamma_{\{\Phi_{put}, V_{\widehat{\sigma}}\}} = \Gamma_{\{\Phi_{call}, V_{\widehat{\sigma}}\}}$ *and* $\Gamma_{\{\Phi_{put}, V_{\widetilde{\sigma}}\}} = \Gamma_{\{\Phi_{call}, V_{\widetilde{\sigma}}\}}$*, where by* $f_{N(0,1)}$ *we denote the probability density function of a standard normal random variable and*

$$d_1 = \frac{\log(v/K) + \left(r + \frac{\widehat{\sigma}^2}{2}\right)(T-t_0)}{\widehat{\sigma}\,\sqrt{T-t_0}}\,,$$
$$\widetilde{d}_1 = \frac{\log(v/K) + \left(r + \frac{L^2\,\sigma^2}{2}\right)(T-t_0)}{L\,\sigma\,\sqrt{T-t_0}}\,.$$

Proof. By Remark 9.4.3 we have that $\partial^2_{SS}\Pi(t,S,B) = \partial^2_{VV}\Pi(t,V)\,\varphi(t)^2$. Let us compute $\partial^2_{VV}\Pi(t,V)$:

$$\begin{aligned} \partial^2_{VV}\Pi(t,V) &= \partial_V\left[\partial_V\Pi\left(t_0, \Phi_{\text{call}}(V_{\widehat{\sigma}}(T))\right)\right] \\ &= \partial_V N(d_1) \\ &= N'(d_1)\,\partial_V d_1 \\ &= f_{N(0,1)}(d_1)\,\frac{1}{\widehat{\sigma}\,v\,\sqrt{T-t_0}}\,. \end{aligned}$$

Then, since $\varphi(t_0) = \frac{v\,\widehat{\sigma}}{s\,\sigma}$, we obtain (9.4.12).

The Gamma for put VT options is the same as the Gamma for call VT options, since the second partial derivatives with respect to the portfolio value of the price of the two options are identical. □

In Figure 9.4 we compared the Gamma for standard European options with the Gamma for European options written on VT portfolios. Notice that, while the Gamma for standard European options exhibits two asymptotes only when the underlying risky asset is ATMF, i.e. $S = K\,e^{-r\,(T-t_0)}$ which implies

$$\Gamma_S = \frac{1}{s\,\sigma\,\sqrt{T-t_0}} e^{-\frac{1}{2}\frac{\sigma^2\,(T-t_0)}{4}}\,,$$

for the Gamma of VT options we have always two asymptotes, since for low volatilities also their Gamma is amplified, in fact even more than the Delta.

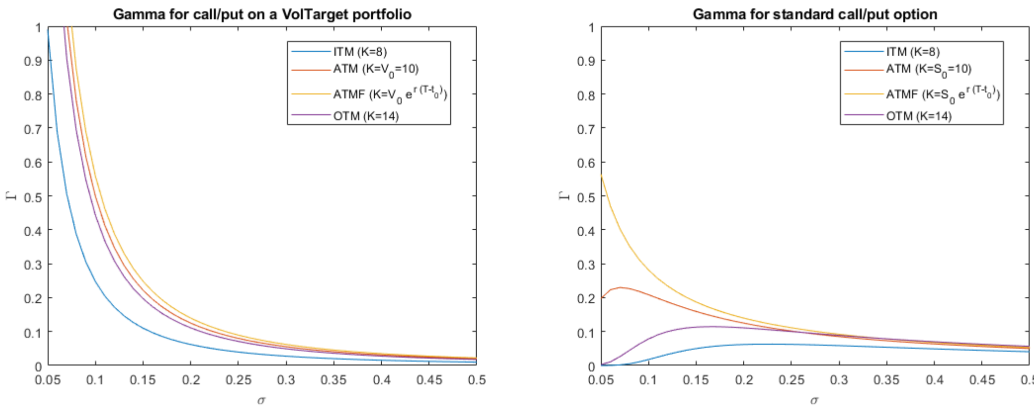

Figure 9.4: The right graph represents the behavior of the Gamma of standard call/put options with respect to different volatility values. The left graph represents the Gamma for call/put options written on VolTaget portfolios. As in Figure 9.3, we considered strike prices K in order to obtain an "in-the-money" option, an "at-the-money" option, an "at-the-money-forward" option and an "out-of-the-money" option. Notice that, once again, also this Greek for the VT options exhibits asymptotes, while this is the case for standard options only when the underlying asset is ATMF. The parameters are fixed as $v = s = 10$, $\widehat{\sigma} = 0.2$, $\mu = 8\%$, $r = 5\%$, $T = 1$, $t_0 = 0$ and the volatilities start at $\sigma = 0.05$.

9.5 Numerical simulations

To better explain how a VolTarget portfolio works, let us assume that the dynamics of the risky asset evolve according to the Heston model, see e.g. [8, 64, 84] for further details on the model, i.e.

$$\mathrm{d}S_t = \mu\, S_t\, \mathrm{d}t + \sqrt{\nu_t}\, S_t\, \mathrm{d}W_t^{(1)}, \tag{9.5.1}$$

$$\mathrm{d}\nu_t = \kappa\,(\theta - \nu_t)\, \mathrm{d}t + \xi\, \sqrt{\nu_t}\, \mathrm{d}W_t^{(2)}, \tag{9.5.2}$$

where $W^{(1)}$ and $W^{(2)}$ are two Brownian motions with correlation parameter ρ, ν evolves as a CIR process and represents the instantaneous variance of the risky asset, θ is the long-variance, κ the rate at which ν reverts to θ, ξ is the volatility of the volatility, and we assume that the Feller condition holds:

$$2\,\kappa\,\theta > \xi^2,$$

in order to guarantee the process ν to be strictly positive.

Let us consider underlying risky asset's parameters calibrated to values observed in the real data, as in the papers by Morellec et al. [93], and CIR parameters as in the seminal paper by Samuelson [107]. We consider the parameters as shown in Table 9.1. which are an adaptation of the ones in [107, 93] in order to show representative scenarios explaining the effect of the VolTarget and the VolTarget maximum leverage strategies.

Fig.	κ	θ	ξ	ρ	ν_0	μ	S_0	r	B_0	V_0	T
9.5	0.6067	0.2207	0.2928	-0.75	0.2154	8.24%	100	4.2%	20	100	1
9.6	"	0.1707	"	"	0.1654	"	"	"	"	"	"

Table 9.1: The parameters of the Heston model that we considered in order to simulate the asset dynamics.

We partition the time-interval $[0, T]$ into N equal subintervals of width T/N

$$0 = t_0 < t_1 < \cdots < t_N = T$$

and simulate a realization of the bivariate process $(S, \nu)_t$. Then, through the following modified *Euler-Maruyama scheme*, we approximate the path of the corresponding VolTarget portfolio

$$V_{\widehat{\sigma}}^{\Delta t}(t_{n+1}) = V_{\widehat{\sigma}}^{\Delta t}(t_n)\left\{1 + \frac{\alpha(t_n)}{S(t_n)}\Delta S_n + \frac{1-\alpha(t_n)}{B(t_n)}\,\Delta B_n\right\}, \quad \text{for } n \in \{0, \dots, N-1\}, \tag{9.5.3}$$

where we defined $\Delta S_n := S(t_{n+1}) - S(t_n)$, $\Delta B_n := B(t_{n+1}) - B(t_n)$ and $V^{\Delta t}(0) = v$.

Proposition 9.5.1 *Let $T > 0$ be a fixed constant. The numerical scheme (9.5.3) is strongly convergent to the solution to (9.1.4), i.e.*

$$\lim_{\Delta t \to 0} \mathbb{E}[|V_T - V_T^{\Delta t}|] = 0\,.$$

Proof. The Euler-Maruyama scheme associated to equation (9.1.4) is

$$V_{\widehat{\sigma}}^{\Delta t}(t_{n+1}) = V_{\widehat{\sigma}}^{\Delta t}(t_n)\left\{1 + \left[\frac{\widehat{\sigma}}{\sqrt{\nu(t_n)}}(\mu - r) - r\right]\Delta t + \widehat{\sigma}\,\Delta W_n^{(1)}\right\}, \tag{9.5.4}$$

which i strongly convergent to (9.1.4). Moreover, we have that

$$\lim_{\Delta t \to 0} \mathbb{E}\left[\frac{\frac{\Delta S_n}{S(t_n)} - \mu\,\Delta t}{\sqrt{\nu(t_n)}} - \Delta W_n^{(1)}\right] = 0\,. \tag{9.5.5}$$

Substituting (9.5.5) in (9.5.4), we obtain (9.5.3). □

The Euler-Maruyama scheme is strongly convergent, but only with order 0.5, and therefore, since the converging order is poor, it is more convenient to consider the following modification of *Milstein scheme*:

$$V_{\widehat{\sigma}}^{\Delta t}(t_{n+1}) = V_{\widehat{\sigma}}^{\Delta t}(t_n) \left\{ 1 + \alpha(t_n) \frac{\Delta S_n}{S(t_n)} + (1 - \alpha(t_n)) \frac{\Delta B_n}{B(t_n)} \right. \\ \left. - \frac{\alpha(t_n)\,(1 - \alpha(t_n))}{2} \left[\left(\frac{\Delta S_n}{S(t_n)} \right)^2 - \nu(t_n)\,\Delta t \right] \right\},$$

which in general converges strongly to the solution with order 1.

9.5.1 Performance analysis

One can notice that the sensitivity analysis of the previous section is widely shown in Figure 9.6. For example the tendency of the path is more or less met, dependently on the volatility instantaneous value; see Section 9.4.2 for the treatment of the sensitivity of the VT portfolio with respect to small changes in the underlying risky asset (the Delta). Moreover, it is clearly visible that the white noise is affecting the VT portfolio value linearly, i.e. the Vega for the VT standard portfolio is null, see Section 9.4.1.

9.6 Extension to the transaction case and concluding remarks

In the present chapter we have presented one of the first attempts to consider options linked to VolTarget strategies from an analytical perspective. We have developed closed end formulas for call and put options linked to VolTarget concepts. Furthermore, we have derived also for selected Greeks closed-end formulas.

The results are in line with what we would expect from a practitioner point of view. One can see, how a VolTarget approach can simplify option pricing for structured products and why also key hedging parameters look much easier than for standard options with changing volatility pattern. Of course we made some simplifying assumptions to obtain the derived results. We aim at developing more general framework within forthcoming researches as to relax (at least some of) the aforementioned assumptions.

As to make an example, we aim at looking how to incorporate transaction costs within our framework. In particular, we are going to consider two alternatives to reach complete such a task. The first one consists in a modification of the chosen VolTarget strategy, i.e. the VolTarget portfolio will no longer pursue a constant volatility, instead it will aim to have a volatility belonging to a desired interval. The second possibility consists in a structural modification, in which the times of portfolio weight adjustments will be restricted to a discrete subset. In other words, the VolTarget strategy will pursue the target volatility only in these discrete times, instead of continuous adjustments. Such modifications are required, when the asset dynamics is assumed to not have a constant volatility, to avoid the cumulated transaction costs to be theoretically infinity, even when the transaction costs are relatively small. These new settings are clearly more challenging from a Mathematical point of view, but also widely treated in the literature, see, e.g. [26, 27, 60, 85, 88], and significantly more appealing from a practitioner point of view.

We also aim at investigating how dynamic asset allocation strategies can be concretely developed within *real world* scenarios. Such a topic is of particular importance from the practitioners' point of view since portfolio weight re-balancing on continuous time base cannot be done, because of the discrete time nature of all financial operations in real markets.

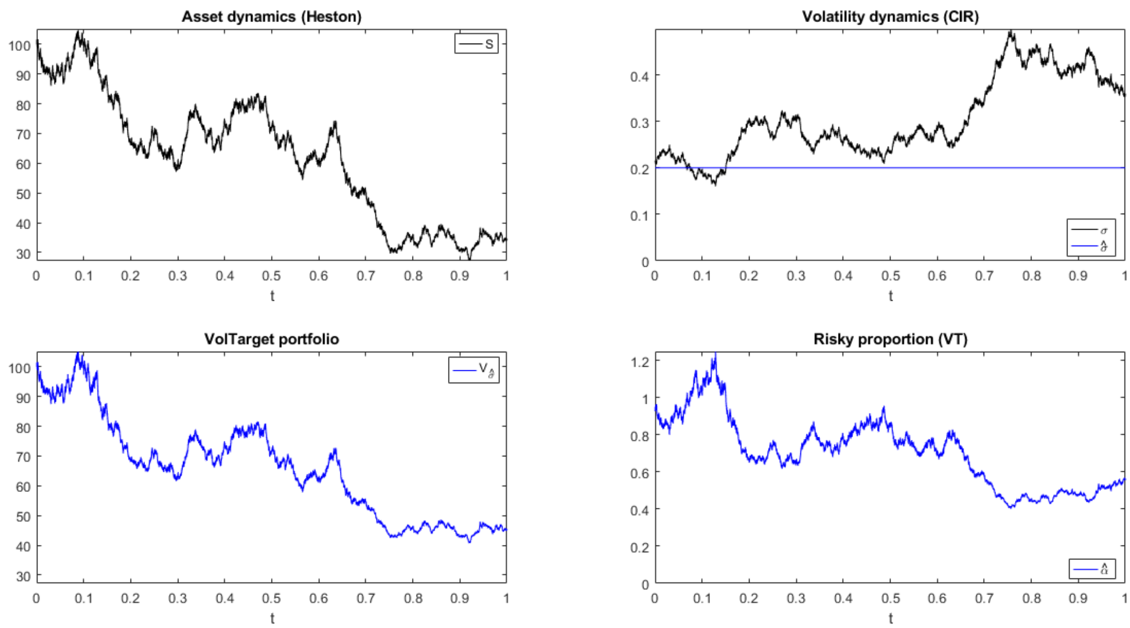

Figure 9.5: The top graphs represent the asset price and the volatility values, simulated as a realization of a Heston model. Here the volatility (top-right figure) is more frequently greater than the target volatility $\widehat{\sigma} = 0.2$, and in these cases the risky proportion $\widehat{\alpha}$ is less than one (bottom-right figure). In the bottom-left figure is represented the corresponding realization of the VolTarget portfolio. For the discretization scheme (9.5.3) we considered a time step $\Delta t = 10^{-6}$.

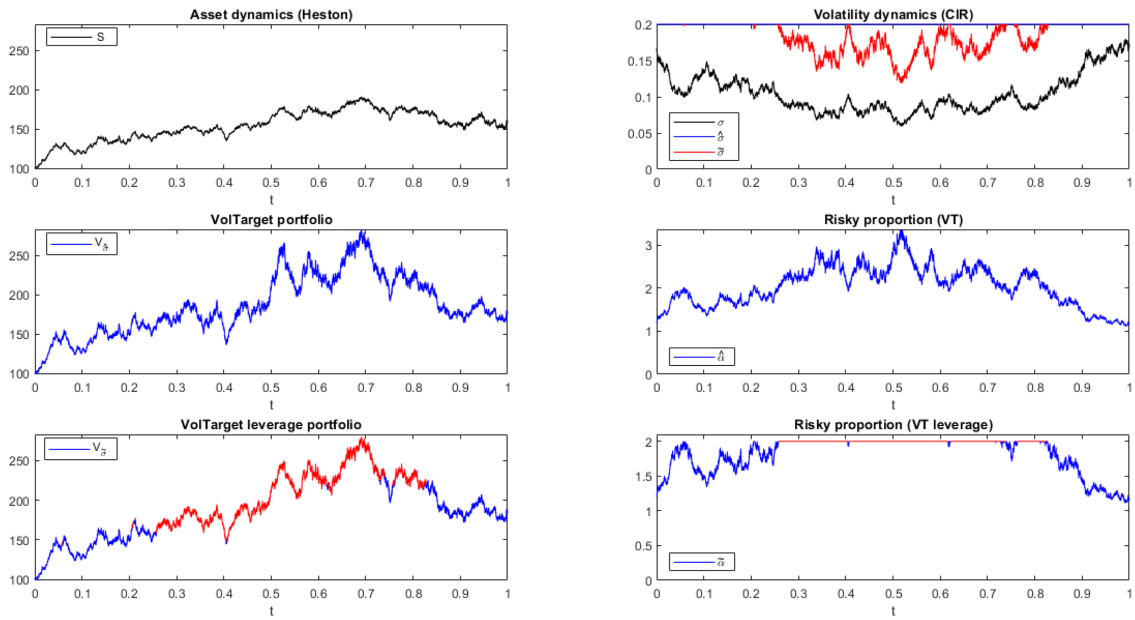

Figure 9.6: The left graphs from the top to the bottom represent the risky asset dynamics, the VT portfolio and the VT leverage portfolio, for $\widehat{\sigma} = 0.2$ and $L = 2$. The top-right figure represents the volatility of the risky asset (in black), the volatility of the VT portfolio (in blue) and the effect of the leverage limitation in the VT strategy (in red). In the bottom-left figure are highlighted in red the path section of the VT leverage portfolio in which the leverage effect intervenes. For the discretization scheme (9.5.3) we considered a time step $\Delta t = 10^{-6}$.

Bibliography

[1] Aase K., *Optimal portfolio diversification in a general continuous-time model*, Stochastic Processes and Applications, 18, 81-98 (1984).

[2] Acerbi C., Tasche D., *On the coherence of the Expected Shortfall*. J. Bank. Financ. 26(7), 1487-1503 (2002).

[3] Acerbi C., Szekely B., *Backtesting Expected Shortfall*. MSCI Research paper (2014).

[4] Aït-Sahalia Y., Cacho-Diaz J., Hurd T.R., *Portfolio Choice with Jumps: a Closed-Form Solution*, The Annals of Applied Probability, 19(2), 556-584 (2009).

[5] Aït-Sahalia Y., Hurd T.R., *Portfolio Choice in Markets with Contagion*, Journal of Financial Econometrics, Volume 14, Issue 1, Pages 128 (2016).

[6] Albeverio S., Steblovskaya V., Wallbaum K., *Investment Instruments with Volatility Target Mechanism*, Quantitative Finance, Vol. 13, No. 10, 1519-1528 (2013).

[7] Albeverio S., Steblovskaya V., Wallbaum K., *The Volatility Target Effect in Structured Products with Protection*, Review of Derivatives Research (2017).

[8] Barbu V., Benazzoli C., Di Persio L., *Feedback optimal controllers for the Heston model*, Applied Mathematics & Optimization (2018).

[9] Basel Committee on Banking Supervision, *Amendment to the Capital Accord to Incorporate Market Risks*, Paper 119 (1996).

[10] Basel Committee on Banking Supervision, *International Convergence of Capital Measurement and Capital Standards*, Paper 128 (2006).

[11] Basel Committee on Banking Supervision, *Guidelines for computing capital for incremental risk in the trading book.*, Paper 159 (2009).

[12] Basel Committee on Banking Supervision,*Revisions to the Basel II market risk framework - final version*, Paper 158 (2009).

[13] Basel Committee on Banking Supervision, *Minimum capital requirements for market risk*, Paper 352 (2016).

[14] Bates T., Kahle M., Stulz R., *Why do U.S. firms hold so much more cash than they used to?* Journal of Finance 64, 1985-2021 (2009).

[15] Benazzoli C., Di Persio L., *Default contagion in financial networks*, Int. Journ. of Math. and Comp. in Simulation, Vol. 10, pp.112-117 (2016).

[16] Benson R., Furbush T., Goolgasian C., *Targeting Volatility: A Tail Risk Solution When Investors Behave Badly*, The Journal of Index Investing, 4 (4), 88-101 (2014).

[17] Bielecki T.R., Rutkowski M., *Credit risk: modeling, valuation and hedging*, Springer Science & Business Media (2013).

[18] Björk T., *Arbitrage Theory in Continuous Time*, Oxford Finace Series, Third edition, ISBN: 9780199574742 (2009).

[19] Black F., Cox J.C., *Valuing corporate securities: Some effects of bond indenture provisions*, The Journal of Finance 31.2: 351-367 (1976).

[20] Black F., Scholes M., *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, 3, 659-683 (1973).

[21] Blinder J.J., Merges J.M., *Stock Market Volatility and Economic Factors*, Review of Quantitative Finance and Accounting 17 (1), 5-26 (2001).

[22] Bokanowski O., Picarelli A., Zidani H., *State-constrained stochastic optimal control problems via reachability approach*, SIAM Journal on Control and Optimization 54.5: 2568-2593 (2016).

[23] Bonollo M., Di Persio L., Mammi L., Oliva I., *Estimating the Counterparty Risk Exposure by Using the Brownian Motion Local Time*, Int. Journ. of Appl. Math.and Comp. Science, 27, (2), pp. 435-447 (2017).

[24] Borodin A.N., Salminen P., *Handbook of Brownian Motion - Facets and Formulae*, Springer Basel AG, second edition (1996).

[25] Bottai M., Cai B., McKeown R.E., *Logistic quantile regression for bounded outcomes. Statistics in Medicine*, Stat. Med. 29 (2), pp. 309-317 (2010).

[26] Boyle P.E., Vorst T., *Option replication in discrete time with transaction costs*, Journal of Finance 47 (1), 271 (1992).

[27] Boyle P.P., Emanuel D., *Discretely adjusted option hedges*, Journal of Financial Economics 8, 259l (1980).

[28] Bremaud P., *Point Processes and Queues: Martingale Dynamics*, Springer (1981).

[29] Bouchard B., Elie R., Touzi N., *Stochastic target problems with controlled loss*, SIAM Journal on Control and Optimization 48.5: 3123-3150 (2009).

[30] Bouchard B., Elie R., Imbert C., *Optimal control under stochastic target constraints*, SIAM Journal on Control and Optimization 48.5: 3501-3531 (2010).

[31] Bouchard B., Dang N.M., *Generalized stochastic target problems for pricing and partial hedging under loss constraints-application in optimal book liquidation*, Finance and Stochastics 17.1: 31-72 (2013).

[32] Campello M., Graham J., Harvey C., *The Real Effects of Financial Constraints: Evidence from a Financial Crisis*, Journal of Financial Economics 97, 470-487 (2010).

[33] Capponi A., Chen P.C., *Systemic risk mitigation in financial networks*, Available at SSRN 2293426 (2013).

[34] Chen E.J., *Two-phase quantile estimation*. IEEE Xplore 1, pp. 447-455. DOI: 10.1109/WSC.2002.1172916 (2002).

[35] Chevalier E., Ly Vath V., Scotti S., *An Optimal Dividend and Investment Control Problem under Debt Constraints*, SIAM Journal on Financial Mathematics, 4:1, 297-326, (2013).

[36] Cordoni F., Di Persio L., *A maximum principle for a stochastic control problem with multiple random terminal times*, arXiv preprint arXiv:1801.07216 (2018).

[37] Cordoni F., Di Persio L., *A BSDE with delayed generator approach to pricing under counterparty risk and collateralization*, International Journal of Stochastic Analysis 2016 (2016).

[38] Cordoni F., Di Persio L., *Gaussian estimates on networks with dynamic stochastic boundary conditions*, Infinite Dimensional Analysis, Quantum Probability and Related Topics 20.01: 1750001 (2017).

[39] Cordoni F., Di Persio L., *Stochastic reaction-diffusion equations on networks with dynamic time-delayed boundary conditions*, Journal of Mathematical Analysis and Applications 451.1: 583-603 (2017).

[40] Costantinides G.M., *Capital Market Equilibrum with Transaction Costs*, Journal of Political Economy, 94, 842-862 (1986).

[41] Crandall M., Ishii H., Lions P.L., *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., **27**, 1-67 (1992).

[42] Dai M., Zhang Q., Zhu Q., *Trend Following Trading under a Regime Switching Model*, SIAM Journal on Financial Mathematics 1, 780-810 (2010).

[43] David H.A., Nagaraja H.N., *Order Statistics*. Wiley Ser. Probab. Stat (2005).

[44] Davis M.H.A., Norman A.R., *Portfolio selections with transaction costs*, Mathematics of Operations Research, Vol. 15, No. 4 (1990).

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

[46] Davis M.H.A., Zariphopoulou T., *American options and transaction fees*, In *Mathematical Finance* M. Davis et al Eds., Springer-Verlag, New York, 47-61 (1995).

[47] Decamps J.P., Villeneuve S., *Optimal dividend policy and growth option*, Finance and Stochastics, 11, 3-27 (2007).

[48] Diamond D.W., *Reputation acquisition in debt markets*, J.Polit. Economy 97 828-862 (1989).

[49] Dixit A., Pindick R., *Investment under uncertainty*, Princeton University Press (1994).

[50] Docherty P., Wang G., *Using synthetic data to evaluate the impact of RTGS on systemic risk in the Australian payments system*, J. Financ. Stab.6 (2), 103-117 (2010).

[51] Duchin R., Ozbas O., Sensoy B., *Costly External Finance, Corporate Investment, and the Mortgage Credit Crisis*, Journal of Financial Economics 97, 418-435 (2010).

[52] Duckworth K., Zervos M., *A model for investment decisions with switching costs*, Annals of Applied Probability, 11, 239-250 (2001).

[53] Duffie D., Sun T.S., *Transaction Costs and Portfolio choice in a Discrete-Continuous Time Setting*, Journal of Economic Dynamics and Control (1990).

[54] Eaton J., Gersovitz M., *Debt with Potential Repudiation: Theoretical and Empirical Analysis*, The Review of Economic Studies, 48(2), 289-309 (1981).

[55] Eisenberg L., Thomas H.N., *Systemic risk in financial systems*, Management Science 47.2: 236-249 (2001).

[56] Escobar M., Kiechle A., Seco L., Zagst R., *Option on a CPPI*, International Mathematical Forum, Vol. 6, 2011, no. 5, 229 - 262 (2011).

[57] European Parliament, *Regulation (EU) No 575/2013 on prudential requirements for credit institutions and investment firms and amending Regulation* (2013).

[58] Framstad N.C., Øksendal B., Sulem A., *Optimal consumption and portfolio in a jump diffusion market with proportional transaction costs*, Journal of Mathematical Economics, 35, 233-257 (2001).

[59] Glasserman P., *Risk horizon and rebalancing horizon in portfolio risk measurement*. Math. Finance, 22(2), pp. 215249 (2012).

[60] Gilster J.E., *The systematic risk of discretely rebalanced option hedges*. Journal of Financial and Quantitative Analysis 25 (4), 507 (1990).

[61] Gordy M.B., *A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules*. J. Financial Intermediation, 12 (2003).

[62] Graham J., Harvey C., *The theory and practice of corporate finance: evidence from the field*, J. Finan. Econ. 60 187-243 (2001).

[63] Grüne L., Picarelli A., *Zubovs method for controlled diffusions with state constraints*, Nonlinear Differential Equations and Applications NoDEA 22.6: 1765-1799 (2015).

[64] Heston S.L., *A Closed-Form Solution for options with Stochastic Volatility with Applications to Bond and Currency options*, The Review of Financial Studies. 6 (2): 327343 (1993).

[65] Hilliard J., Li W., *Volatilities implied by price changes in the S& P 500 options and futures contracts*, Review of Quantitative Finance and Accounting 42 (4), 599-626 (2013).

[66] Ho L., Cadle J., Theobald M., *An analysis of risk-based asset allocation and portfolio insurance strategies*, Review of Quantitative Finance and Accounting 36 (2), 247-267 (2011).

[67] Hugonnier J., Malamud S., Morellec E., *Capital Supply Uncertainty*, Cash Holdings, and Investment. Review of Financial Studies , 28, 391-445 (2015).

[68] Hugonnier J., Malamud S., Morellec E., *Credit market frictions and capital structure dynamics*, Journal of Economic Theory, 157, 1130-1158 (2015).

[69] Jawaid H., *Pricing and hedging of the European option linked to target volatility portfolio*, SSRN Electronic Journal. http://ssrn.com/abstract=2743699. doi: 10.2139/ssrn.2743699 (2015).

[70] Jawaid H., *An analysis of guaranteed lifetime withdrawal benefits linked to target volatility portfolio*, SSRN Electronic Journal. http://ssrn.com/abstract=2743780. doi: 10.2139/ssrn.2743780 (2016).

[71] Jeanblanc M., Shiryaev A., *Optimization of the flow of dividends*, Russian Math. Survey, 50, 257-277, (1995).

[72] Kabanov Y., Sararian M., *Financial Markets with Transaction Costs*, Springer (2010).

[73] Kitchin J., *Cycles and Trends in economic factors*, Review of Economics and Statistics. Vol 5, pp 10-16, (1923).

[74] Knuth D., *The art of computer programming*: Volume 3. Sorting and Searching, Second Edition. Addison-Wesley. ISBN 0-201-89685-0 (1988).

[75] Koenker R., Bassett G., *Regression quantiles*, Econometrica, 46 (1), pp. 33–50 (1978).

[76] Kou S.G., *A Jump-Diffusion Model for Option Pricing*, Management Science, 48(8), 1086-1101 (2002).

[77] Kraft H., *Optimal portfolios and Heston stochastic volatility model: an explicit solution for power utility*, Quantitative Finance, 5(3), 303-313 (2007).

[78] Krein D., Fernandez J., *Indexing Volatility Risk Control*, The Journal of Index Investing, 3 (2), 62-75 (2012).

[79] Kuhn C., Stroh M., *Optimal portfolios of a small investor in a limit order market: a shadow price approach*, Math. Finan. Econ., 3, 45-72 (2010).

[80] Hodges S., Neuberger A., *Optimal replication of contingent claims under transaction costs*, Review of Futures Markets, 8, 222-239 (1989).

[81] Jeanblanc M., Pontier M., *Optimal portfolio for a small investor in a market model with discontinuous prices*, Applied Mathematics and Optimization, 22(1), 287-310 (1990).

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

[83] Laurenta J.P., Sestiera M., Thomasb S., *Trading book and credit risk: how fundamental is the Basel review*, J. Bank. Financ., 73 (1) (2016).

[84] Lee C., Sokolinskiy O., *R-2GAM stochastic volatility model: flexibility and calibration*, Review of Quantitative Finance and Accounting 45 (3), 463-483 (2015).

[85] Leland H.E., *Option pricing and replication with transaction costs*, Journal of Finance 40 (5), 1283 (1985).

[86] Lipton A., *Modern monetary circuit theory, stability of interconnected banking network, and balance sheet optimization for individual banks*, International Journal of Theoretical and Applied Finance 19.06: 1650034 (2016).

[87] Manna M., Nobili S., *Banks holdings of and trading in government bonds*, Bank of Italy Working paper, n.1166 (2018).

[88] Mello A.S., Neuhaus H.J., *A portfolio approach to risk reduction in discretely rebalanced option hedges*, Management Science 44 (7), 921 (1998).

[89] Merton R.C., *Lifetime portfolio selection under uncertainty: The continuous-time case*, Rev. Econom. Statist. **51** 247-257 (1969).

[90] Merton R.C., *Optimum consumption and portfolio rules in a continuous time model*, J. Economic Theory **3** 373-413 (1971).

[91] Merton R.C., *On the pricing of corporate debt: The risk structure of interest rates*. The Journal of finance 29.2: 449-470 (1974).

[92] Modigliani F., Miller M., *The cost of capital, corporation finance and the theory of investment*, American Economic Review 48, 261-297 (1958).

[93] Morellec E., Nikolov B., Schurhoff N., *Corporate Governance and Capital Structure Dynamics*, Journal of Finance, Vol. 67, No. 3, pp. 803-848. http://dx.doi.org/10.1111/j.1540-6261.2012.01735.x (2012).

[94] Moscadelli M., *The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee*. Bank of Italy working paper, n. 517 (2004).

[95] Mugnolo D., *Semigroup methods for evolution equations on networks*, Springer (2014).

[96] MSCI RiskMetrics Group, *CreditMetrics technical document* (2007).

[97] Ozgoren M.K., Longman R.W., Cooper C.A., *Probabilistic inequality constraints in stochastic optimal control theory*, Journal of Mathematical Analysis and Applications 66.1: 237-259 (1978).

[98] Page L., Brin S., Motwani R., Winograd T., *The PageRank citation ranking: Bringing order to the web*, Technical Report. Stanford InfoLab (1999).

[99] Patton A., Sheppard K., *Optimal combination of realised volatility estimators*, Intern. J. Forecasting, Vol. 25(2), 218-238 (2009).

[100] Pham H., *Stochastic control under progressive enlargement of filtrations and applications to multiple defaults risk management*, Stochastic Process. Appl. 120, no. 9, 1795-1820 (2010).

[101] Pham H., *On the smooth-fit property for one-dimensional optimal switching problem*, Sminaire de Probabilits, vol XL 187-201 (2007).

[102] Predescu M., Wilkens S., *Default risk charge: modeling framework for the Basel risk measure*, J. Risk, 19:4 (2017).

[103] Protter P.E., *Stochastic Integration and Differential Equations*, Springer, Second Edition, Version 2.1 (2005).

[104] Revuz D., Yor M., *Continuous Martingales and Brownian Motion*, Springer, Third Edition (2005).

[105] Revyakin G., *A new approach to the nature of business cycles and their analysis in the global context*, Eureka: Social and Humanities. 5: 27-37 (2017).

[106] Rogers L.C.G., Veraart L.A.M., *Failure and rescue in an interbank network*, Management Science 59.4: 882-898 (2013).

[107] Samuelson P.A., *Mathematics of Speculative Price*, SIAM Review Vol. 15, No. 1, pp. 1-42 (1973).

[108] Siao J.H., Hwang R.C., Chu C.K., *Predicting recovery rates using logistic quantile regression with bounded outcomes*, Quant. Finance, 16(5) pp. 777-792 (2016).

[109] Shreve S., Soner M., *Optimal investment and consumption with transaction costs*, The Annals of Applied Probability, Vol. 4, No. 3, 609-692 (1994).

[110] Shreve S.E., *Stochastic calculus for Finance II : continuous-time models*, New York; London: Springer (2004).

[111] Taksar M., Klass M.J, Assaf D., *Optimal investment and consumption with transaction costs*, The Annals of Applied Probability, Vol. 4, No. 3, 609-692 (1994).

[112] Øksendal B., Sulem A., *A Diffusion Model for Optimal Portfolio Selection in the Presence of Brokerage Fees*, Mathematics of Operation Research, 13, 277-294 (1988).

[113] Øksendal B., A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Berlin Heidelberg New York: Springer (2007).

[114] Soner H.M., Touzi N., *Dynamic programming for stochastic target problems and geometric flows*, Journal of the European Mathematical Society 4.3: 201-236 (2002).

[115] Vasicek O., *The Distribution of Loan Portfolio Value*, Risk (2002).

[116] Villeneuve S., Warin X., *Optimal liquidity management and hedging in the presence of a non-predictable investment opportunity*, Math. Financ. Econ. 8 193-227 (2014).

[117] Zakamulin V., *Dynamic Asset Allocation Strategies Based on Unexpected Volatility*, Journal of Alternative Investments, Vol. 16, No. 4, 2014. Available at SSRN: http://ssrn.com/abstract=2357881or http://dx.doi.org/10.2139/ssrn.2357881 (2013).

[118] Zeng X., Taksar M., *A stochastic volatility model and optimal portfolio selection*, Quantitative Finance, 13(10), 1547 (2016).

[119] Zervos M., Johnson T., Alazemi F., *Buy-low and sell-high investment strategies*, Math. Finance 23:3, 560-578 (2013).