Utility Indifference Pricing of Credit Instruments

by

Georg Sigloch

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto

Copyright © 2009 by **Georg Sigloch**

Abstract

Utility Indifference Pricing of Credit Instruments

Georg Sigloch

Doctor of Philosophy

Graduate Department of Mathematics

University of Toronto

2009

While the market for credit instruments grew continuously in the decade before 2008, its liquidity has dried up significantly in the current crisis, and investors have become aware of the possible consequences of being exposed to credit risk. In this thesis we address these issues by pricing credit instruments using utility indifference pricing, a method that takes into account the investor's personal risk aversion and which is not affected by the lack of liquidity.

Through stochastic optimal control methods, we use indifference pricing with exponential utility to determine corporate bond prices and CDS spreads. In the first part we examine how these quantities are affected by risk aversion under different models of default. The emphasis lies on a hybrid model, in which a regime switch of the reference entity is triggered by a creditworthiness index correlated to its stock price.

The second part generalizes this setup by introducing uncertainty in the model parameters. Robust optimal control has been used independently in the literature to address model uncertainty for portfolio selection problems. Here, we incorporate this approach with utility indifference and derive some analytical and numerical results on how model uncertainty affects credit spreads.

Acknowledgements

Foremost I would like to express my deep gratitude to my supervisor Professor Sebastian Jaimungal from the Department of Statistics at the University of Toronto for helping me develop the topic of this thesis, sharing his ideas and supporting me throughout my work.

Furthermore I would like to thank Professor Luis Seco and Professor Jeremy Quastel from the Department of Mathematics, as well as Professor Marcel Rindisbacher, formerly at Rotman School of Management (now at Boston University School of Management), for acting as members on my Ph.D. committee and for providing valuable suggestions for my research. Moreover, I strongly appreciate the thorough reading of my thesis as well as the various comments made by my external examiner, Professor Tom Salisbury from the Department of Mathematics and Statistics at York University.

Finally I am thankful to the graduate administrator of the Department of Mathematics, Ida Bulat, for always being extremely helpful and patient when I had questions regarding organizational matters.

Contents

0 Introduction **1**

0.1 Credit Risk and Credit Securities 1

0.2 Pricing Theory 6

0.3 Utility Indifference Pricing 9

0.4 Portfolio Optimization 13

0.5 Outline and Contributions of this Thesis 18

1 Preliminaries **23**

1.1 The Merton Investment Problem with Exponential Utility 23

1.2 The Merton Problem with a Defaultable Stock 27

1.A Appendix: The Inverse of a Symmetric Matrix 36

2 Indifference Pricing in a Reduced Form Model **38**

2.1 Without Investment in the Defaultable Stock 39

2.2 With Investment in the Defaultable Stock 44

2.A Appendix: Verification Theorems 49

3 Indifference Pricing in a Structural Model of Default **53**

3.1 Introduction 53

3.2 The Model 55

3.3 The Investment Problem 56

3.4 The Defaultable Bond 60
3.5 The Credit Default Swap 64
3.A Appendix: Verification Theorem 66
3.B Appendix: The Heat Equation on the Quarter Plane 70

4 Indifference Pricing with Invisible Default Boundary 73

4.1 The Model 75
4.2 The Investment Problem 78
4.3 Reduction of Dimension 85
4.4 Pricing of Credit Derivatives 87
4.A Appendix: Verification Theorems 90

5 Indifference Pricing under Model Misspecification 95

5.1 Motivation 95
5.2 The Investment Problem 99
5.3 Valuation of Credit Derivatives 110
5.A Appendix: Verification Theorem 115
5.B Appendix: Motivation for the Definition of $\mathbf{\Phi}$ 121

6 Pricing of a CDS in a Discrete Time Setting 124

6.1 Introduction and Model 124
6.2 The Two Investment Problems 126
6.3 Computation of U 128
6.4 Behaviour of the Indifference Price as $\gamma \to 0$ 129
6.A Appendix: Proof of Theorem 20 134

7 Conclusions and Directions for Future Work 140

7.1 Conclusions 140
7.2 Future Work 141

Bibliography **144**

Chapter 0

Introduction

0.1 Credit Risk and Credit Securities

0.1.1 Products

While fixed-income instruments like corporate bonds or loans have existed for a long time, credit derivatives are fairly young. Many of them were introduced in the early 1990s and can be viewed as a second generation of derivatives (after derivatives on equities and commodities): while the latter were first introduced to manage market risk, credit derivatives were invented to manage and hedge credit risk. Originally tailor-made OTC solutions, their number and volume has dramatically increased over the years. Consequently, the market for many credit derivatives used to be fairly liquid (until 2008). This fact also made them attractive to investors for the pure purpose of speculation.

In the literature there are several similar, but slightly different versions for the definition of a *credit derivative*. Since in this thesis we are only working with specific examples, and a precise general definition is irrelevant, the following definition (taken from Schönbucher (2003)) seems to be appropriate:

Definition 1. *(a) A credit derivative is a derivative security that is primarily used to transfer, hedge or manage credit risk.*

(b) A credit derivative is a derivative security whose payoff is materially affected by credit risk.

The most common underlying assets for credit derivatives are loans and their securitized versions, bonds. A bond, loan or mortgage is a contract between two counterparties. At the time of entry into the contract the creditor lends money to the obligor, for which the latter agrees to pay back a predetermined amount (the *face value* or *notional*) at maturity. In the case of a zero-coupon bond, these are the only payments agreed upon, while for a coupon-bearing bond, the obligor makes additional periodic predetermined *coupon payments*.

If the creditor enters this kind of contract, he is exposed to *credit risk*, namely the risk of losing his investment in the case a *credit event* occurs. A credit event is defined as the obligor's default, i.e. the failure to meet his obligations. Possible credit events include bankruptcy, failure to make coupon payments, restructuring or downgrade below a certain level.

When a credit event occurs, the assets of the bond issuer are normally liquidated to meet his obligations at least partially. Consequently, bond holders can expect to receive a certain percentage of the notional even in the case of a default. This percentage is called *recovery rate* and ideally paid at or very shortly after default. In reality however, the settlement process can take quite a long time.

A credit derivative is a contract between two counterparties A and B. Party A pays B a predetermined fee and in return receives payments from party B which depend on the occurrence or non-occurrence of a credit event of a third party C between now and the maturity of the credit derivative. This third party C is also known as the *reference entity* or *reference credit*. If the payments depend only on the default of a single reference entity, the credit derivative is called *single-name*; if there are several reference entities, it is called *multi-name*.

The most common single-name credit derivative is the credit default swap, even

though there are also multi-name CDSs. Its original purpose was to provide insurance for the holder of a corporate bond against default of the reference entity. The buyer of a CDS makes periodic or continuous payments from the time of entry into the contract until maturity or time of default of the reference entity – whichever occurs first. If there is no default until maturity, the seller of the CDS does not have to make any payments. If default occurs before maturity, he is obliged to buy the bond his counterparty at face value at the time of default. In the case of instant recovery payment, this is equivalent of making a payment of $(1 - R)F$ to the buyer of protection, where R is the recovery rate of the bond and F its face value. Credit default swaps are the type of credit derivatives on which we focus in this thesis.

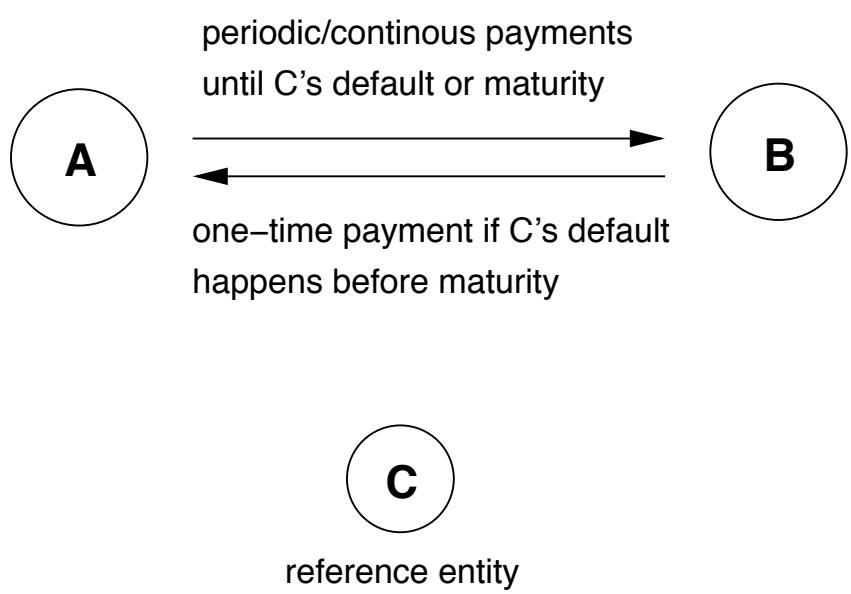


Figure 1: Counterparties in a CDS contract

Other common single-name credit derivatives are asset swaps and total return swaps, while the most common multi-name credit derivatives are collateralized debt obligations (CDOs). A CDO can viewed as an insurance of credit securities from a whole basket. The credit risk is bundled and redistributed to investors in several tranches. Since this thesis is not concerned with any of the instruments mentioned above, we refer to Hull

(a) Continuous payments

(b) Discrete time payments

Figure 2: Cash flows for a CDS from the seller's point of view. F is the face value, R the recovery rate and A the CDS spread.

(2005) or Schönbucher (2003) for details.

When this thesis was started (2005), credit derivatives were a very fast growing market, which peaked in 2007/2008. According to Giesecke (2009), industry sources estimated the notional of credit derivatives outstanding at 62 trillion USD. As it is well-known, especially CDSs and CDOs made a significant contribution to the current financial crisis (2008/2009). As it became clear that uncontrolled redistribution of credit risk poses a imminent danger to the worldwide financial system, the outstanding notional of credit derivatives had been reduced to under 20 trillion USD by late 2008 (also according to Giesecke (2009)). Consequently, the liquidity for many credit derivatives has dried up.

This fact certainly raises the question whether the credit derivatives market (specifically credit default swaps) still has a future and whether it is still worth putting effort into their pricing. Since the original purpose of CDSs was to hedge credit risk, and since there will still be need for this in the future, it is safe to say that both questions can be answered with "yes". However it is also almost certain that products will be held simple and will be subject to more regulation than in the past. Furthermore, the market for credit derivatives will probably not be as liquid as it used to be close to its peak. However, as it will be explained in section 0.3, utility indifference pricing is not affected by this fact and may even gain importance for this reason in the context of credit derivative pricing.

0.1.2 Models

When pricing credit derivatives or underlying assets, one first has to choose an underlying model of the reference entity's default. In the literature, mainly three different types of models have been used: *reduced form* models, *structural* models and *hybrid* models.

In reduced form models (also called *intensity based* models), default is triggered by an exogeneous event independent of the internal structure of the reference entity. The most common approach (see e.g. Jarrow and Turnbull (1995)) is to model the *hazard rate process* as a Poisson process with a given deterministic hazard rate, whose switching triggers default. This approach typically has the advantage that the model parameters can be calibrated fairly easily. In the traditional risk-neutral pricing, so called *risk-neutral* hazard rates are extracted from observed market prices of other traded, related credit derivatives (concerning the idea of risk-neutral pricing see the next section). The emphasis here is on *traded* – while there used to be a liquid market for many credit instruments, especially bonds and credit default swaps, many of these securities have become illiquid as a consequence of the recent financial crisis. However, even if parameters can be calibrated, experience shows that default of a firm is always at least somewhat correlated to its internal structure, which makes reduced form models unrealistic.

In contrast, the structural approach models default as a consequence of a company becoming *unhealthy*. A popular structural model is the *firm value* model, which measures the company's health by its firm value, which itself is the viewed as the sum of the company's equity and debt. This interpretation makes the firm value a traded asset. Default occurs when the value of the firm is less than its outstanding debts or some percentage of the outstanding debt. In the first paper using a structural approach, Merton (1974) models a company's equity as a European call option on its firm value with its debts used as a strike level. The company defaults if at maturity the value of the firm is below the company's debts. The advantage of this model is its simplicity even though it lacks realism. Black and Cox (1976) extend this idea to the more realistic case

where the company defaults the instant its firm value drops below a critical level, turning the problem into a first passage time one.

There have been numerous extensions, modifications and increases in the sophistication of the firm value model over the last several decades. A limited list of important contributions to the field include Leland (1977) extension of the debt to a coupon paying bond; Kim, Ramaswamy, and Sundaresan (1993) and Longstaff and Schwartz (1995) inclusion of stochastic interest rates; Leland (1994) and Leland and Toft (1996) extension to endogenously specifying the default boundary as a result of equity holders maximizing the value of the firm; Duffie and Lando (2001) model of incomplete market information; and Fouque, Sircar, and Solna (2006) integration of stochastic volatility in firm value models.

Hybrid models combine the underlying ideas of reduced-form and structural models. Here default is triggered by the switching of a Cox process, i.e. a counting process with *stochastic* hazard rate. The hazard rate is normally assumed to be negatively correlated to the firm value – it increases, when the firm value decreases, and the other way round.

0.2 Pricing Theory

The fundamental principle in pricing theory assumes that in an ideal financial market, there are no arbitrage opportunities. In real world, arbitrage opportunities do exist – but only for very short time periods. In the theory of derivative pricing we mainly distinguish between complete and incomplete markets. A market is called *complete*, if every claim can be replicated perfectly, i.e. at time 0 the investor can set up a portfolio and has an adapted trading strategy which replicates the payoff of the claim perfectly at maturity. In a complete market under the absence of arbitrage, the price of any claim is uniquely determined as the value of its replicating portfolio. Simple examples of complete markets (see e.g. Björk (2004)) include the one-period binomial model with

a money market account and a risky asset. An easy continuous-time example is the standard Black-Scholes model with a money market account and a risky stock modeled as an Ito diffusion.

When a claim can be replicated perfectly, there are two main methods of finding its value. Firstly, one can determine the corresponding replicating strategy, i.e. the number of units of tradable assets that are needed at any time between 0 and maturity to replicate the claim . Using the replicating strategy at time 0 and the prices of the tradable assets, one then easily computes the value of the replicating portfolio.

Secondly, an investor who is only interested in pricing a claim, but not in replicating it, can omit the latter and apply the *fundamental theorem of asset pricing* (FTAP) instead. One of the statements of the FTAP is that under the absence of arbitrage, there exists a measure Q, which is equivalent to the real-world measure $\mathbb{P}$, and under which the *discounted* price processes of all tradable assets are martingales. As a consequence, the value of the value of the claim can be computed as

$$C_0 = \mathbb{E}^Q[e^{-rT} C_T], \tag{1}$$

where, for simplicity, we have assumed constant interest rates. More generally, the price of C at time $t \in [0, T]$ is given by

$$C_t = \mathbb{E}^Q[e^{-r(T-t)} C_T \mid \mathcal{F}_t].$$

In a complete market, the *equivalent martingale measure* Q is unique, and hence the price of the claim is uniquely determined by the formula above, as it should be by the replication argument.

It should be pointed out that the assumptions made in the common complete market models are very strict. For example, one normally has to assume that the market is frictionless (i.e. no transaction costs), that trading is possible continuously (at least in continuous time models), that borrowing money is possible at the same rate as lending, etc. It is therefore safe to say that in real world, markets involving exclusively big cap

stocks are almost complete, but not exactly, while markets involving small cap stocks and derivatives are substantially incomplete.

Another prime example of an incomplete market is one that contains credit instruments. By definition, any credit instrument contains a certain credit risk, which usually cannot hedged away by basic assets like stocks or risk-free bonds. The emphasis here is on basic assets – obviously, the default risk of a defaultable bond can be hedged away by other related credit instruments, e.g. an appropriate credit default swap.

In an incomplete market there is no unique method of pricing derivatives. The part of the FTAP quoted above for complete markets still holds, i.e. under no-arbitrage assumptions the value of a claim at time 0 is still given by equation (1). However, the equivalent martingale measure Q is not unique this time, and the investor has to make a decision which measure to choose.

The investor's choice of the appropriate Q is normally a subjective one and depends on his risk preferences. In real the world one can try to determine Q as well as possible from observed market prices. On any underlying asset there is normally more than one claim available. If the dynamics of the asset are given by an Ito diffusion, the market prices of risk of all claims on this asset have to be the same for no arbitrage reasons. Theoretically this enables the investor to determine Q. However, the essential assumption here is that the reference claim is liquidly traded – which is often not the case.

Another way is to choose Q "as close as possible" to the real-world measure $\mathbb{P}$, in the sense that he wants to minimize

$$\mathbb{E}\left[f\left(\frac{dQ}{d\mathbb{P}}\right)\right]$$

over all equivalent martingale measures. Here f is a strictly convex function on $[0,\infty)$. Popular choices are $f(x) = x\ln x$, in which case Q is the *minimal entropy martingale measure* , or $f(x) = x^2$.

Finally, if the seller wants to eliminate any risk for himself, one obvious choice would be to *superreplicate* the claim, which means that he sets up a portfolio which matches

or exceeds the payoff of the claim at maturity almost surely. He then offers to sell the claim at the thereby obtained *upper hedging price* (as in Karatzas and Shreve (1998)). In practice, a potential buyer will often consider this price too high and therefore decline the deal.

A pricing method in incomplete markets that has become increasingly popular over the past years is *utility indifference pricing*. It avoids the problem of identifying the "correct" measure Q altogether. The method uses the real-world measure $\mathbb{P}$ only and does not make any assumptions on continuous tradability of the derivatives to be priced. The main ideas will be presented in the next section.

0.3 Utility Indifference Pricing

Indifference pricing is pricing from the point of view of portfolio optimization. To have an example to work with, let us assume that we would like to price a contingent claim with random payoff C_T, and that the underlying tradable assets are a money market account and a number of risky assets, modeled as stochastic processes on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Let u be a given *utility function*, measuring the performance of the investor. Typically one assumes that that u is strictly increasing, because any "rational" investor prefers having more wealth to having less, and that u is strictly concave, because the investor is risk-averse. When using utility indifference pricing, we compare the two scenarios "do not invest in the claim" vs. "invest in the claim". For the first scenario, let $W_t^{\pi,w}$ denote the wealth process corresponding to a self-financing strategy of investment in the stock and money market with initial endowment w. Let

$$V(w) \triangleq \sup_{\pi} \mathbb{E}[u(W_T^{\pi,w})] \tag{2}$$

be the corresponding *value function*. In the second scenario, the investor buys one unit of the claim for p dollars and invests his remaining wealth in the money and stock market. If w is his wealth *after* buying the claim, then for a given trading strategy π let $\overline{W}_t^{\pi,w}$ be

the corresponding wealth process, and define

$$\overline{V}(w) \triangleq \sup_{\pi} \mathbb{E}[u(\overline{W}_T^{\pi,w})]. \tag{3}$$

Note that $\overline{W}^{\pi,w}$ differs from $W_t^{\pi,w}$ because of the payoff from the claim (in this example, $\overline{W}_T^{\pi,w} = W_T^{\pi,w} + C_T$). Then the *buyer's indifference price* of the claim is the value p which satisfies

$$\overline{V}(w-p) = V(w), \tag{4}$$

i.e. the buyer is indifferent between being invested and not invested in the claim. Similarly, one can define the seller's indifference price by replacing $w-p$ by $w+p$ in the equation above. It is important to note that the indifference price is a personal price, and therefore perhaps the term indifference *value* seems to be more appropriate. Not only does it differ for the buyer and the seller, but it also depends on the investor's level of risk aversion. Only if the market price is lower (higher) than this, will the buyer (seller) engage in the transaction.

The utility-based approach of evaluating an investor's performance was suggested by Von Neumann and Morgenstern (1944), who showed that under certain assumptions on their behaviour, investors try to maximize expected utility according to their risk preferences. For a detailed introduction and justification, see also Föllmer and Schied (2002). Indifference pricing was introduced by Hodges and Neuberger (1989). For this reason, in the literature the indifference price is sometimes called *Hodges price*.

When trying to find the indifference price, the main difficulty normally lies in solving the two portfolio optimization problems (2) and (3). The standard methods are dynamic programming and martingale methods, which will be explained in section 0.4. Note however that often a numerical solution of these problems is not good enough, since one has to find the solution of (3) in dependence of the parameter p, and then solve equation (4) for p. This naturally limits the choice of utility functions. Furthermore, to keep the optimization problems simple enough, one often has to assume constant parameters in

the dynamics of the assets.

While one does not have to identify an appropriate risk-neutral measure to apply indifference pricing, it should be mentioned that one has an equally difficult problem, namely determining the appropriate investor's utility function reflecting his personal risk aversion. Nevertheless, the determination of this utility function does *not* depend on the tradability of the given assets. This thesis however is not concerned with this problem. Throughout this thesis, we will always assume that u is given as *exponential utility*:

Assumption 1.

$$u(x) = -\frac{1}{\gamma}e^{-\gamma x}$$

for some $\gamma > 0$.

The parameter γ measures the risk aversion of the investor: $\gamma = 0$ corresponds to complete risk-neutrality (for which the optimization problems above normally do not have a finite solution), whereas a large γ corresponds to a high level of risk aversion. In the literature exponential utility is most frequently used. The main reason for this fact is mathematical tractability: as it will be seen, exponential utility allows one to factor out wealth from the value function, thereby reducing the dimension of the optimization problem at hand.

Important quantities to classify utility functions are the coefficients of *absolute risk aversion*, defined as $-\frac{u''(x)}{u'(x)}$, and the coefficient of *relative risk aversion*, given by $-x\frac{u''(x)}{u'(x)}$. It is easy to check that exponential utility has constant absolute risk aversion, whereas power utility $u(x) = \frac{x^p}{p}$, $-\infty < p < 1$, has constant relative risk aversion. According to Cochrane (2001), the latter is more realistic, and consequently, it would be desirable to use power utility instead. This however would be at the cost of analytical tractability.

In their original paper Hodges and Neuberger (1989) demonstrate how transaction costs can be analyzed in this framework. Since its introduction, a large number of papers have been published in utility indifference pricing. Davis, Panas, and Zariphopoulou (1993) study the impact on derivative pricing in the presence of transaction costs.

Some more recent contributions include Henderson and Hobson (2002a) who explore options on a non-tradable asset correlated to a tradable one, and Stoikov (2005) investigates the implications for volatility derivatives. Ilhan, Jonsson, and Sircar (2004) price options in incomplete markets and study the impact of stochastic volatility on the prices.

Two other standard examples of incomplete markets are the those of insurance/reinsurance contracts and weather/energy derivatives. Jaimungal and Nayak (2004) and Young and Zariphopoulou (2002) examine the former, while Carmona (2008a) contains work on the latter.

For the reason mentioned earlier, most papers use exponential utility. One of the few that uses a different utility function is Henderson and Hobson (2002a), where the pricing of contingent claims under power utility is examined. Due to the inseparability of wealth, the authors have to make a few approximation assumptions and obtain their results as asymptotic expansions. Brendle and Carmona (2004) work with exponential, power and logarithmic utility and at the same time generalize their results to partially observable markets.

In the context of using indifference pricing to price credit instruments, only little work seems to have been done. In Sircar and Zariphopoulou (2007), the authors study single and two-name credit derivatives in reduced-form and hybrid models. In Sircar and Zariphopoulou (2009) they generalize this setup for the pricing of CDOs. The paper most closely related to the material covered in this thesis is Leung, Sircar, and Zariphopoulou (2008), where corporate bonds in a structural framework are considered. More details how our work relates and distinguishes itself from this paper can be found in section 0.5.

Finally, it should be mentioned that a good introduction and overview of the subject is Henderson and Hobson (2004). Several of the classical papers listed above have been collected and published in Carmona (2008b).

0.4 Portfolio Optimization

In this section we explain two different methods of solving utility-based portfolio optimization problems, as they frequently appear in indifference pricing problems. The methods presented are martingale methods and dynamic programming. In this thesis only dynamic programming will be used, because this method seems to be most appropriate. Occasionally martingale methods are easier to use, and in the literature there are several papers where this is done.

To see how both methods work, let us consider one of the simplest problems: given a risk-free money market account and n risky assets (stocks) $S^{(1)}, \ldots, S^{(n)}$ which are modeled as Ito diffusions on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, the investor's goal is to

$$\text{maximize} \quad \mathbb{E}[u(W_T)], \quad \text{given that} \quad W_0 = w,$$

over all possible trading strategies. Here the utility function u is assumed to be strictly increasing and concave, $W_T^{\boldsymbol{\pi}}$ is the wealth process corresponding to a given trading strategy $\boldsymbol{\pi}$, and $T > 0$ is given. We require that trading strategies be $\mathcal{F}_t$-adapted. A trading strategy can be characterized as a vector-valued process $\boldsymbol{\pi}_t \in \mathbb{R}^n$, where $\pi_t^{(i)}$ is the dollar amount invested in $S^{(i)}$ at time t, and the remaining wealth is invested in the money market. We only consider *self-financing* wealth processes, i.e. any change in wealth results from changes in the asset prices, but not from additional money coming from external sources.

0.4.1 Martingale Methods

References for Martingale methods are e.g. in Karatzas, Lehoczky, Shreve, and Xu (1991), Karatzas and Shreve (1998), Korn (1997). In its basic version, the martingale method is only applicable if the market at hand is complete. We therefore assume that

the dynamics of the tradable assets are given by

$$dM_t = r_t M_t \, dt,$$
$$d\boldsymbol{S}_t = \mathrm{diag}(\boldsymbol{S}_t) \; [\boldsymbol{\mu}_t \, dt + \boldsymbol{\sigma}_t \, d\boldsymbol{B}_t] \,,$$

where $\boldsymbol{S}_t = (S_t^{(1)}, \; \ldots, \; S_t^{(n)})$, and $r_t \in \mathbb{R}$, $\mu_t \in \mathbb{R}^n$, $\boldsymbol{\sigma}_t \in \mathbb{R}^{n\times n}$ are $\mathcal{F}_t$-adapted, and $\boldsymbol{B}_t$ is an n-dimensional standard Brownian motion on Ω. Then we model the wealth process corresponding to a trading strategy $\boldsymbol{\pi}$ by the dynamics

$$dW_t = \left[rW_t + (\boldsymbol{\mu}_t - \boldsymbol{r}_t)^T \boldsymbol{\pi}_t\right] \; dt + \boldsymbol{\pi}_t \boldsymbol{\sigma}_t d\boldsymbol{B}_t,$$
$$W_0 = w,$$

with $\boldsymbol{r}_t = (r_t, \; \ldots, r_t)^T \in \mathbb{R}^n$.

Since the market is complete is complete, there exists a unique measure Q which is equivalent to $\mathbb{P}$ and under which the discounted stock price processes $\tilde{S}_t$ are martingales (here discounted always means discounted by M_t) . Let η_t denote the Radon-Nikodym derivative of the corresponding measure change. It is straightforward to show that under Q the discounted wealth process $\widetilde{W}_t$ is a martingale as well, which implies that

$$w = W_0 = \widetilde{W}_0 = \mathbb{E}^Q[\widetilde{W}_t] = \mathbb{E}^{\mathbb{P}}[\eta_T M_T^{-1} W_T^{\boldsymbol{\pi}}].$$

Hence the problem can be rewritten as

$$\text{maximize} \quad \mathbb{E}[u(W_T^{\boldsymbol{\pi}})] \quad \text{subject to} \quad \mathbb{E}[\eta_T M_T^{-1} W_T^{\boldsymbol{\pi}}] = w.$$

This constraint problem can be solve using a generalized version of the Lagrange multipliers method. The completeness of the market then guarantees the existence of a trading strategy such that the maximum from above can be attained. Note however that is not easy to determine this trading strategy explicitly.

In the references Karatzas, Lehoczky, Shreve, and Xu (1991), Karatzas and Shreve (1998), Korn (1997) it is also shown how this method can be generalized to incomplete markets.

0.4.2 Dynamic Programming

In presentation of this section we follow Øksendal (2007). Standard references on dynamic programming and stochastic optimal control in general are Fleming and Rishel (1975), Fleming and Soner (2005) and Yong and Zhou (1999) .

When using the dynamic programming method, we do not have to assume that the given market is complete. However we do have to assume that the coefficients of the given assets are Markov, and we also have to restrict ourselves to Markov trading strategies. It is applicable to many problems from stochastic optimal control, not only portfolio optimization. Pioneering work includes Bellman (1952) and Kushner (1962), among others.

We assume that the state process of a controlled stochastic system on Ω is given by the controlled stochastic equation

$$dX_t = dX_t^{\boldsymbol{\pi}} = \boldsymbol{b}(t, X_t, \boldsymbol{\pi}_t)\ dt + \boldsymbol{\sigma}(t, X_t, \boldsymbol{\pi}_t)\ d\boldsymbol{B}_t.$$
$$X_0 = x_0.$$

$\boldsymbol{B}_t$ is a Brownian motion on Ω, and the control $\boldsymbol{\pi}_t$ is a stochastic process adapted with respect to the given filtration.

The goal is to maximize

$$\mathbb{E}[u(X_t)]. \tag{5}$$

To invoke the principle of dynamic programming, we introduce the time and state dependent *performance function*

$$J^{\boldsymbol{\pi}}(x,t) \triangleq \mathbb{E}[u(X_t) \mid X_t = x\ ]$$

and define the *value function* as

$$V(x,t) \triangleq \sup_{\boldsymbol{\pi}} J^{\boldsymbol{\pi}}(x,t).$$

Then (5) is given by $V(x_0, 0)$. The fact that even though we are only interested in finding (5) for one initial condition of X_0, we have to solve a similar optimization problem for all possible states of (X_t, t), is characteristic for dynamic programming.

The key to finding V is *Bellman's principle of optimality.* In our example, it states that $V(X_t^{\boldsymbol{\pi}}, t)$ is a supermartingale for any control $\boldsymbol{\pi}$, and if an optimal control $\boldsymbol{\pi}^*$ exists, then $V(X_t^{\boldsymbol{\pi}^*})$ is a martingale. This can be summarized as

$$V(x,t) = \sup_{\boldsymbol{\pi}} \mathbb{E}[V(X_{t+h}^{\boldsymbol{\pi}}, t+h) \mid X_t^{\boldsymbol{\pi}} = x\ ]$$

for all $(x,t) \in [0,T] \times \mathbb{R}^n$ and all h such that $0 \leq h \leq T - t$.

Assuming that V is regular enough, we can apply Ito's lemma to the right hand side, divide by h and let $h \to 0$ to get a partial differential equation that V has to satisfy, namely

$$\partial_t V + \sup_{\boldsymbol{\pi} \in \mathbb{R}^n} \mathcal{L}^{\boldsymbol{\pi}} V = 0, \tag{6}$$

where

$$\mathcal{L}^{\boldsymbol{\pi}} V(x,t) = \sum b_i(t,x,\boldsymbol{\pi})\ \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum a_{ij}(t,x,\boldsymbol{\pi})\ \frac{\partial^2 V}{\partial x_i \partial x_j}, \quad (a_{ij}) = \boldsymbol{\sigma}^T \boldsymbol{\sigma}.$$

Equation (6) is the so-called *Hamilton-Jacobi-Bellman (HJB)* equation. From the definition of V we are also given a boundary condition at $t = T$, namely

$$V(x,T) = u(x) \quad \text{for all} \quad x \in \mathbb{R}. \tag{7}$$

It is normally hard to prove from its definition, that V is sufficiently regular. In fact, there exist easy-looking examples of optimization problems whose value functions are not differentiable everywhere (see e.g. Øksendal (2007), p. 248–250). However, in many cases there exist verification theorems of the following type:

If there exists a function H which is continuous and $C^{(1,2)}$ on $[0,T] \times \mathbb{R}^n$ which satisfies (6), (7) and which is sufficiently integrable, then $H = V$ for $(x,t) \in [0,T] \times \mathbb{R}^n$.

In this thesis we mainly consider optimization problems for which solutions of the HJB equation can be found, and therefore verification theorems of this type appear frequently.

The term "sufficiently integrable" normally depends on the problem to be solved and will be described on a case to case basis.

The main method to solve the HJB equation, if possible, is the following: one finds the optimal control $\boldsymbol{\pi}^*$ in terms of the partial derivatives of V. In the context of portfolio optimization problems, $\mathcal{L}^{\boldsymbol{\pi}}V$ is often quadratic in $\boldsymbol{\pi}$, and hence this step is straightforward. One then substitutes this $\boldsymbol{\pi}^*$ back into (6). The result is a highly non-linear partial differential equation. It should therefore be pointed out, that in most cases, it is not possible to solve the HJB equation in closed form. However if it can be solved, we can get the optimal strategy from the representation of $\boldsymbol{\pi}^*$ from the first step above, which also applies if we get a numerical solution for U.

There are numerous variations and extensions of the stochastic optimization problem covered here. Firstly, it is straightforward to define the value function not only in terms of terminal wealth but also by consumption over time. The consumption process then acts as a second control, apart from the trading strategy. Furthermore, the method is not limited to cases when the state process is a diffusion. For example, Øksendal and Sulem (2007) is a reference for optimal control of jump diffusions. Bensoussan (1992) generalizes the setup to the situation when not all state variables are observable to the controller.

In the context of portfolio optimization, the dynamic programming principle was first used in Merton (1969). Since then, a lot of extensions and generalizations have been published. To name a few, Davis and Norman (1990) include transaction costs into their analysis, Korn (1999) introduce impulse controls, and Uppal and Wang (2003) include model uncertainty from a robust investor's point of view.

0.5 Outline and Contributions of this Thesis

In this section we summarize the content of the chapters following this introduction and explain their contributions. As explained in section 0.3, the main tool in indifference pricing is portfolio optimization. Besides introducing some more or less standard notation, we examine two portfolio optimization problems in chapter 1. The first of these problems is an investment problem which was first treated in Merton (1969) and can be found in many standard textbooks. In this thesis we will call it the *standard Merton investment problem*, and we recall it for the reader's convenience. It considers a market with a money market account and a number of risky assets modeled by geometric Brownian motions. The investor's goal is to maximize expected utility of his terminal wealth.

The second investment problem is a slight variation of the first, and it seems (to G. Sigloch's best knowledge) not to be covered in the literature. The market still consists of a money market account and a number of risky assets, but one of these assets is assumed to be defaultable. Default is triggered by the switching of a Poisson process independent of the driving Brownian motions. Since default cannot be anticipated in this setup, the investor loses the money invested in the defaultable stock. We show that the resulting optimization problem leads to an ordinary differential equation which cannot be solved analytically, but fairly easily by using numerical methods.

Chapter 2 uses utility indifference pricing to price corporate bonds and credit default swaps in a simple reduced form model. We consider a market with a money market account, one risky, default-free asset (e.g. a stock index) and a defaultable asset, which is assumed to be the reference entity's stock. As in the second part of chapter 2, default is triggered by the switching of an independent Poisson process. We consider two different scenarios: in the first case, the investor is not allowed to invest in the defaultable stock, while in the second case, the investor does not have this restriction. However if the stock defaults, he loses the money invested in it. The results obtained for bond yields and CDS spreads show significant differences. This is a clear confirmation, that the

indifference price of credit instruments (or derivatives in general) does not only depend on the investors risk aversion, but also strongly on the available tradable assets in which he is allowed to invest.

There are certainly more sophisticated and complicated reduced-form models than the one described above. Since the focus of this thesis is on structural models of default, we do not pursue them here. Sircar and Zariphopoulou (2007) is one example where the authors analyze the effect of risk aversion of single and two-name cases within a reduced form approach, and more general, in a hybrid model. In Sircar and Zariphopoulou (2009) the same authors apply indifference pricing to CDOs, also in reduced-form and hybrid models. However, both these papers lack the realism that upon default the investor loses the money he had invested in the firm's stock.

Chapter 3 addresses indifference pricing in a simple structural framework. The original version of this chapter was Jaimungal and Sigloch (2008), a paper which was submitted and accepted for presentation for the Bachelier Congress 2008. However, we withdrew the paper, since only a very short time before, Leung, Sircar, and Zariphopoulou (2008) introduced a very similar setup. In this paper, the authors consider a market model with a money market account and a defaultable risky asset, and use utility indifference pricing to price defaultable bonds on this risky asset. In that work, the firm's stock price and its asset value are modeled as correlated geometric Brownian motions. However, in contrast to previous models, although the asset value is assumed observable, it is not tradable.

We extend the model in Leung, Sircar, and Zariphopoulou (2008) in several aspects. Firstly, there is no reason to assume that the defaultable stock is the only available tradable asset. A real world investor is always able to invest in many liquid stocks, and more importantly, investors will try to diversify their portfolios. As a consequence, we consider a market in which the investor also trades in a correlated non-defaultable index. This setting can easily be extended to several default-free risky assets, but we will not do this here. Secondly, experience shows that it is not reasonable to assume

that default of a company can be completely anticipated. Consequently, we assume that after a *credit worthiness index* (CWI; comparable to the firm value in Leung, Sircar, and Zariphopoulou (2008)) crosses a certain threshold D, the state of the company changes from *healthy* to *distressed*. At this point the company does not default, and instead enters a state of financial distress, in which default is now triggered by an exogenous Poisson process. In this context, D can be interpreted as a rough upper estimate of an otherwise unknown default barrier, after whose hitting investors become nervous and withdraw their investments from the firm. Another interpretation of D is that of the level at which rating agencies downgrade the credit rating of the company. The model presented is therefore not a pure structural model, but a hybrid model in the sense that we have two different regimes for the state of the reference entity.

The methods to solve the highly non-linear HJB equations in this chapter were first introduced by Zariphopoulou (2001) and used in the context of substitute hedging in Henderson and Hobson (2002b). These papers however only consider the case with one tradable asset, and it is interesting to see that the method can be generalized to the case of several tradable assets.

Chapters 4 and 5 form the heart of this thesis. The model introduced in chapter 3 as well as the simple structural model in Black and Cox (1976) both have the unsatisfying property that credit spreads tend to 0 for short maturities. The reason for this behavour is quite obvious: Since both the CWI (or the firm value) and the critical barrier D are observable, the firm's survival can be anticipated for short times to maturity, if the CWI is above D at the present point in time. In reality however, non-zero credit spreads are observed even for short maturities. To explain non-zero credit spreads for short maturities in structural models, Duffie and Lando (2001) and Giesecke (2006) introduced models where the investor is only given partial information on model parameters. One possibility presented in the latter paper, and which we will adopt in chapter 4, is to assume that the critical barrier D is invisible, i.e. modeled as a time-invariant random variable. The

information given to the investor is then given by the paths of the tradable assets and the CWI up to the current time, as well as whether or not default has occurred yet – however, if the firm is still alive, he *cannot* observe how far away the barrier D still is.

Very interestingly, the method for solving the HJB equations from chapter 3 still works in this setup, and it turns out that after simplifying, we have to find a function which solves a system of coupled heat equations. This solution has to be found numerically, which is subject to future research.

In chapter 5 we address the very real fact that some model parameters may be uncertain. In particular, this concerns the CWI since usually the perceived health of a company can only be fully observed a few times a year, e.g. when the firm publishes its earnings. While it would be desirable to introduce the CWI as an unobservable quantity, in the indiffernce pricing setting this would lead to a highly non-tractable problem. Instead, we take a different approach.

We follow ideas from Anderson, Hansen, and Sargent (2000), Maenhout (2004) and Uppal and Wang (2003), who introduce model uncertainty to portfolio optimization problems. We assume that the investor has a rough estimate $\mathbb{P}$ for the real-world measure, but due to model uncertainty he is also willing to consider alternative equivalent measures. We adapt methods from robust portfolio optimization and augment the optimization problem to incorporate a minimax problem where one maximizes expected *penalized* utility of terminal wealth over all admissible trading strategies while minimizing over a set of measures equivalent to the historical one. The penalty as together with the minimization act as a control on how far from the original measure the investor is willing to deviate. Even though the HJB equations are significantly more complicated due to the underlying min-max problem, we derive closed form solutions. Moreover, we examine how the bond yields and CDS rates behave under this model as uncertainty increases/decreases.

The contents of chapter 6 was the project which started the research for this thesis. In

a model similar to the one in Merton (1974), we use indifference pricing to price a credit default swap. It is different and almost independent from the previous chapters, because it assumes discrete-time payments and discrete-time monitoring of default, while in previous chapters both was done in continuous time. The methodology is very similar as before and should be no surprise at this point. However the value function of the optimization problem corresponding to an investment in the CDS is now defined piecewise, due to the discrete monitoring. The main part is to derive an explicit expression for this value function and to compare the resulting indifference CDS spread to the risk-neutral spread.

Chapter 7 finally concludes this thesis. It summarizes the results and gives an overview over possible directions for future research.

Chapter 1

Preliminaries

In this chapter we introduce some basic concepts and notation in portfolio optimization. In section 1.1 we review the well-known Merton optimization problem with n tradable risky, but default-free assets, which can be found in many standard textbooks. In section 1.2 we discuss a variation in which one of the stocks is defaultable.

1.1 The Merton Investment Problem with Exponential Utility

We examine the scenario in which there are n tradable risky assets $S^{(1)}, \ldots, S^{(n)}$, as well as a risk-free asset M_t (the money market account). The risky assets are modeled as correlated geometric Brownian motions with constant coefficients, i.e.

$$dS_t^{(i)} = S_t^{(i)} \left(\mu_i \, dt + \sigma_i \, dB_t^{(i)}\right), \quad i = 1, \ldots, n.$$

The covariance matrix of $S^{(1)}, \ldots, S^{(n)}$ is denoted by $\mathbf{\Omega}$. The underlying filtered probability space is denoted by $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. In this basic setup we typically assume that

$$\mathcal{F}_t = \sigma\left(\{B_s^{(1)}, \ \ldots, \ B_s^{(n)} : \ 0 \leq s \leq t\} \cup \mathcal{N}\right),$$

where $\mathcal{N}$ is the collection of subsets of $\mathbf{\Omega}$ of measure 0. The filtration $(\mathcal{F}_t)$ satisfies the *usual conditions*: it is right-continuous, increasing, and $\mathcal{F}_0$ contains all the sets of measure 0.

Let an $\mathcal{F}_t$-adapted process $\boldsymbol{\pi}_t = \left(\pi_t^{(1)},\ \ldots,\ \pi_t^{(n)}\right)$ be given. We interpret $\pi_t^{(i)}$ as the dollar amount invested in $S^{(i)}$ at time t. Then we model the corresponding wealth process $W_t^{\boldsymbol{\pi}}$ corresponding to $\boldsymbol{\pi}$ as the process given by the controlled SDE

$$dW_s = \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_s + r\ W_s\right]\ ds + \sum_{i=1}^{n} \pi_s^{(i)} \sigma_i\ dB_s^{(i)}, \tag{1.1}$$

provided that this SDE has a unique strong solution for any initial condition $W_t^{\boldsymbol{\pi}} = w$. It follows from Yong and Zhou (1999), chapter 6, that this is the case, if the condition

$$\int_0^T \boldsymbol{\pi}_t^2\ dt < \infty \quad \text{a.s.}$$

Moreover, in this case we can represent this solution by the *variation of constant formula*

$$W_s^{\boldsymbol{\pi}} = e^{r(s-t)} \left(w + (\boldsymbol{\mu} - \boldsymbol{r}) \int_t^s e^{-r(\tilde{s}-t)} \boldsymbol{\pi}_{\tilde{s}}\ d\tilde{s} + \sum_{i=1}^{n} \sigma_i \int_t^s e^{-r(\tilde{s}-t)} \pi_{\tilde{s}}^{(i)}\ dB_{\tilde{s}}^{(i)} \right). \tag{1.2}$$

Throughout this thesis we will normally write only W_t instead of $W_t^{\boldsymbol{\pi}}$, if confusion is unlikely.

We define the value function V as

$$V(w, \boldsymbol{S}, t) \triangleq \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}\left[u(W_T) \mid W_t = w,\ \boldsymbol{S}_t = \boldsymbol{S}\right]. \tag{1.3}$$

Here $\boldsymbol{S}_t = (S_t^{(1)},\ \ldots,\ S_t^{(n)})$, $\boldsymbol{S} = (S^{(1)},\ \ldots,\ S^{(n)}) \in \mathbb{R}^n$, and $\mathcal{A}$ is the set of admissible trading strategies to be defined below.

Definition 2. *An $\mathcal{F}_t$-adapted process $\boldsymbol{\pi}_t = \left(\pi_t^{(1)}, \ldots, \pi_t^{(n)}\right)$ is called an admissible trading strategy, if*

(i)

$$\int_0^T \boldsymbol{\pi}_t^2\ dt < \infty \quad a.s.,$$

(ii)

$$\mathbb{E} \int_0^T \boldsymbol{\pi}_t^2\ \left(e^{-\gamma e^{r(T-t)} W_t^{\boldsymbol{\pi}}}\right)^2\ dt < \infty.$$

Due to condition (ii) the process

$$Y_t \triangleq e^{-\gamma e^{r(T-t)} W_t^{\boldsymbol{\pi}}}$$

is an Ito process for all $\boldsymbol{\pi} \in \mathcal{A}$ (i.e. the drift and volatility terms are sufficiently integrable), which has the dynamics

$$dY_t = Y_t \left[-\gamma e^{r(T-t)} \left(\boldsymbol{\mu} - \boldsymbol{r}\right)^T \boldsymbol{\pi}_t + \frac{1}{2}\gamma^2 \left(e^{r(T-t)}\right)^2 \boldsymbol{\pi}_t^T \boldsymbol{\Omega} \boldsymbol{\pi}_t \right] dt + \\ + Y_t \left(-\gamma e^{r(T-t)}\right) \sum_{i=1}^{n} \pi_t^{(i)} \sigma_i \, dB_t^{(i)}.$$

Any trading strategy which, instead of satisfying (i) and (ii), is almost surely bounded, satisfies the conditions in definition 2. This easily follows from equation (1.2). This has two consequences. The optimal trading strategies we obtain for our optimization problems at hand are always almost surely bounded, and hence are admissible in the sense of definition 2 or of the analogous definitions in future chapters. Secondly, one could replace conditions (i) and (ii) by requiring that any admissible strategy be bounded almost surely.

It is straightforward to see that for a given initial condition $W_t = w$, for any admissible trading strategy the wealth process depends on the increments $B_s^{(i)} - B_t^{(i)}$ $(s \geq t)$, but is independent of the initial condition $\boldsymbol{S}$. Therefore, $V = V(w,t)$ is a function of w and t only.

The corresponding HJB equation is

$$\begin{cases} \partial_t V + \sup\limits_{\boldsymbol{\pi} \in \mathbb{R}^n} \mathcal{L}^{\boldsymbol{\pi}} V = 0, \\ \qquad V(w,T) = u(w), \quad w \in \mathbb{R}, \end{cases}$$

where

$$\mathcal{L}^{\boldsymbol{\pi}} V = rw \, \partial_w V + \frac{1}{2} \boldsymbol{\pi}^T \boldsymbol{\Omega} \boldsymbol{\pi} \, \partial_{ww} V + \boldsymbol{\pi}^T (\boldsymbol{\mu} - \boldsymbol{r}) \partial_w V.$$

The maximum is attained at $\boldsymbol{\pi}_t^* = -\frac{\partial_w V}{\partial_{ww} V} \, \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \boldsymbol{r})$ and hence the PDE becomes

$$\partial_t V + rw \partial_w V - \frac{1}{2} (\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \boldsymbol{r}) \frac{\left(\partial_w V\right)^2}{\partial_{ww} V} = 0,$$

subject to the same boundary condition.

This problem has the solution

$$V(w,t) = -\frac{1}{\gamma} e^{a_t w + b_t}, \tag{1.4}$$

where

$$a_t = -\gamma e^{r(T-t)}, \qquad b_t = -\frac{1}{2}(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \boldsymbol{r})(T-t).$$

Throughout this thesis we will stick to these definitions of a_t and b_t. If we denote the market price of risk of $S^{(1)}, \ldots, S^{(n)}$ by Λ, then $\Lambda^2 = (\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \boldsymbol{r})$, and we can rewrite $b_t = -\frac{1}{2}\Lambda^2 (T-t)$. It is noteworthy that the only dependence of V on the dynamics of the risky assets comes in through Λ.

From the computations above it follows that the candidate for an optimal strategy is given by

$$\boldsymbol{\pi}_t^* = \frac{1}{\gamma e^{r(T-t)}} \, \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \boldsymbol{r}).$$

Note that this strategy is constant up to the factor $e^{-r(T-t)}$.

From the definition of V as in (1.3) it is not clear that V is in fact smooth. However, since we have determined a solution of the corresponding HJB equation (1.1) explicitly, we can apply the following verification theorem:

Theorem 2. *If there exists a function $H(w,t)$ which is continuous and $C^{2,1}$ on $\mathbb{R} \times [0,T]$ and solves equation (1.1) and which satisfies*

$$\mathbb{E} \int_0^T \boldsymbol{\pi}_t^2 \left(\partial_w H(W_t^{\boldsymbol{\pi}}, t)\right)^2 \, dt < \infty$$

for all $\boldsymbol{\pi} \in \mathcal{A}$. Moreover, suppose that for all $(w,t) \in \mathbb{R} \times [0,T]$ there exists $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(w,t) \in \mathbb{R}^n$ such that

$$\mathcal{L}^{\boldsymbol{\pi}^*} H = \sup_{\boldsymbol{\pi} \in \mathbb{R}^n} \mathcal{L}^{\boldsymbol{\pi}} H$$

and such that the process defined by $\boldsymbol{\pi}_t^* = \boldsymbol{\pi}^*(W_t, t)$ is an admissible trading strategy. Then $V(w,t) = H(w,t)$ for $(w,t) \in \mathbb{R} \times [0,T]$, and $\boldsymbol{\pi}^*$ is an optimal strategy, i.e. $V(w,t) = \mathbb{E}[u(W_T^{\boldsymbol{\pi}^*}) \mid W_t^{\boldsymbol{\pi}^*} = w]$.

It is easy to check that the solution of the HJB equation as found in equation (1.4) satisfies the conditions in the verification theorem. In particular, since

$$\partial_w V = e^{r(T-t)}\ e^{b_t}\ e^{-\gamma e^{r(T-t)} w},$$

V satisfies the integrability condition for $\partial_w V$ from the verification theorem, due to condition (iii) in definition 2.

1.2 The Merton Problem with a Defaultable Stock

We examine the scenario in which there are n tradable risky assets $S^{(1)}, \ldots, S^{(n)}$, as well as a risk-free asset M_t (the money market account). As before, the risky assets are modeled as correlated geometric Brownian motions with constant coefficients. The difference to the standard Merton problem from the previous section is, that one of the assets, say $S^{(n)}$, is assumed to be defaultable, while the remaining risky assets are default free. Default is modeled by the switching of a Poisson process N_t such that $N_0 = 0$ with intensity κ, which is assumed to be constant. The analysis of the problem can easily be generalized to a deterministic, time-dependent default rate κ_t, but we will not do this here for notational reasons.

Let τ_d denote the default time, i.e. $\tau_d \triangleq \inf\{t \geq 0 :\ N_t = 1\}$.

For $t < \tau_d$, the investor invests in all the risky assets as well as the money market account. Upon default, the value of $S^{(n)}$ drops to 0, and consequently he only invests in the default free assets for $t \geq \tau_d$. Additionally, since default cannot be anticipated, he also loses the money he has invested in $S^{(n)}$ at time t.

Because of the distinctive role of $S^{(n)}$ it is convenient to write the covariance matrix $\mathbf{\Omega}$ of the risky assets in the form

$$\mathbf{\Omega} = \begin{pmatrix} \overline{\mathbf{\Omega}} & \boldsymbol{\omega} \\ \boldsymbol{\omega}^T & \sigma_n^2 \end{pmatrix}.$$

Furthermore, we let $\overline{\boldsymbol{\mu}} = (\mu_1, \ldots, \mu_{n-1})^T$, $\overline{\boldsymbol{r}} = (r, \ldots, r) \in \mathbb{R}^{n-1}$, $\lambda^2 = (\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}})^T \overline{\boldsymbol{\Omega}}^{-1} (\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}})$ be the $(n-1)$-dimensional analogues of $\boldsymbol{\mu}$, $\boldsymbol{r}$ and Λ^2.

Let $\pi_t^{(i)}$ be the dollar amount in $S^{(i)}$ at time t. Then we model the wealth process W_t corresponding to a trading strategy $\boldsymbol{\pi}$ by the dynamics

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ W_t\right]\ dt + \sum_{i=1}^{n} \pi_t^{(i)} \sigma_i\ dB_t^{(i)}, & t < \tau_d, \\ \left[(\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}})^T \boldsymbol{\pi}_t + r\ W_t\right]\ dt + \sum_{i=1}^{n-1} \pi_t^{(i)} \sigma_i\ dB_t^{(i)}, & t > \tau_d, \end{cases}$$

subject to the condition

$$W_{\tau_d} = W_{\tau_d^-} - \pi_{\tau_d}^{(n)}.$$

At any point in time t it is reasonable to assume that the investor has full information on the stock prices $S_s^{(1)}, \ldots, S_s^{(n)}$ for $0 \le s \le t$, but also whether or not $S^{(n)}$ has defaulted yet. Therefore we model the state of information given to the investor as the filtration $(\mathcal{F}_t)$ with

$$\mathcal{F}_t \triangleq \sigma\left(\left\{B_s^{(1)}, \ \ldots, \ B_s^{(n)} : \ 0 \le s \le t\right\} \cup \{t \ge \tau_d\} \cup \mathcal{N}\right).$$

We now define the set of admissible trading strategies, which we denote by $\mathcal{A}$ again for convenience, even though this set is different from the set $\mathcal{A}$ from section 1.1. In the setup of this section it would be wrong to define an admissible strategy as an $\mathcal{F}_t$-adapted process, because in this case at time τ_d, the investor could take into account the default of $S^{(n)}$ when making his investment decision, which is not realistic. We will therefore define an admissible trading strategy as $\mathcal{F}_t$-predictable instead.

Definition 3. *An $\mathcal{F}_t$-predicted process $\boldsymbol{\pi}_t = \left(\pi_t^{(1)}, \ldots, \pi_t^{(n)}\right)$ is called an admissible trading strategy, if*

(i)

$$\pi_t^{(n)} = 0 \quad \textit{for} \quad t > \tau_d,$$

(ii)

$$\int_0^T \boldsymbol{\pi}_t^2\ dt < \infty \quad a.s.,$$

(iii)

$$\mathbb{E}\int_0^T \boldsymbol{\pi}_t^2 \left(e^{-\gamma e^{r(T-t)} W_t^{\boldsymbol{\pi}}}\right)^2 \, dt < \infty.$$

As in the standard Merton investment problem, we define the value function as

$$U(w, \boldsymbol{S}, t) \triangleq \sup_{\boldsymbol{\pi}\in\mathcal{A}} \mathbb{E}\left[u(W_T) \mid W_t = w,\ \boldsymbol{S}_t = \boldsymbol{S},\ t < \tau_d\right],$$

and as before it is easy to see that the wealth process is independent of $\boldsymbol{S}$. Therefore, $U = U(w,t)$ is a function of w and t only.

This time corresponding HJB equation is

$$\begin{cases} \partial_t U + rw\ \partial_w U + \sup\limits_{\boldsymbol{\pi}\in\mathbb{R}^n} \left\{ \frac{1}{2}\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi}\ \partial_{ww} U + \boldsymbol{\pi}^T(\boldsymbol{\mu} - \boldsymbol{r})\ \partial_w U + \right. \\ \qquad\qquad \left. + \kappa \left[V^{(n-1)}(w - \pi_t^{(n)},\ t) - U\right] \right\} = 0, \\ \qquad\qquad U(w,T) = u(w), \quad w \in \mathbb{R}. \end{cases} \tag{1.5}$$

Here $V^{(n-1)}$ is the value function corresponding to an investment in $S^{(1)}, \ldots, S^{(n-1)}$ and the money market account only. The difference to the HJB equation in section 1.1 is the last term on the left hand side, which accounts for a possible switch from the original investment problem to the standard Merton problem with $n-1$ risky assets and a simultaneous loss of $\pi_{\tau_d}^{(n)}$ dollars. From section 1.1 it follows that

$$V^{(n-1)}(w,t) = u(we^{r(T-t)}) \cdot e^{-\frac{1}{2}\lambda^2(T-t)}.$$

In the following we will solve (1.5). As in section 1.1 there is a verification theorem stating that the solution coincides with the value function U. This verification theorem can be found in more general form in appendix 2.A.

To simplify the given ODE, we make an ansatz of the form $U(w,t) = u(we^{r(T-t)})\ g(t)$. Then g satisfies the equation

$$\begin{cases} g' - \kappa g + \inf\limits_{\boldsymbol{\pi}\in\mathbb{R}^n} \left\{ \boldsymbol{\pi}^T(\boldsymbol{\mu} - \boldsymbol{r}) a_t g + \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ a_t^2 g + \kappa e^{-\frac{1}{2}\lambda^2(T-t)} \cdot e^{-a_t \pi_t^{(n)}} \right\} = 0 \\ \qquad\qquad g(T) = 1. \end{cases} \tag{1.6}$$

It is not obvious that this differential equation has a classical solution on $[0, T]$. If such a solution exists, we expect it to be positive, because U is obviously always negative. In the following we will *assume* that (1.6) in fact has a (unique) classical positive solution. We will derive several properties that this solution has to satisfy, and rigorously prove its existence afterwards.

Letting $\widetilde{\boldsymbol{\pi}}_t = a_t \boldsymbol{\pi}_t = -\gamma e^{r(T-t)} \boldsymbol{\pi}_t$, we have to minimize the function

$$f(\widetilde{\boldsymbol{\pi}}) = \widetilde{\pi}^T (\boldsymbol{\mu} - \boldsymbol{r})\ g + \frac{1}{2} \widetilde{\boldsymbol{\pi}}^T \boldsymbol{\Omega} \widetilde{\boldsymbol{\pi}}\ g + \kappa e^{-\frac{1}{2}\lambda^2 (T-t)} \cdot e^{-\widetilde{\pi}^{(n)}}.$$

The first order condition for obtaining the infimum is

$$(\boldsymbol{\mu} - \boldsymbol{r})\ g + \boldsymbol{\Omega} \widetilde{\boldsymbol{\pi}}\ g + \kappa e^{-\frac{1}{2}\lambda^2 (T-t)} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -e^{-\widetilde{\pi}^{(n)}} \end{pmatrix} = \mathbf{0}. \tag{1.7}$$

As long as $g > 0$, f is a convex function in $\boldsymbol{\pi}$, and hence every solution of (1.7) corresponds to a global minimum of f.

To solve this equation, we define

$$\boldsymbol{A} = \begin{pmatrix} \overline{\boldsymbol{\Omega}}^{-1} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix},$$

so that we have

$$\boldsymbol{A\Omega} = \begin{pmatrix} \boldsymbol{I} & \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega} \\ \boldsymbol{\omega}^T & \sigma_n^2 \end{pmatrix}.$$

Multiplying equation (1.7) by $\boldsymbol{A}$ therefore yields

$$\begin{pmatrix} \overline{\boldsymbol{\Omega}}^{-1} (\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}}) \\ \mu_n - r \end{pmatrix} g + \begin{pmatrix} \boldsymbol{I} & \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega} \\ \boldsymbol{\omega}^T & \sigma_n^2 \end{pmatrix} \widetilde{\boldsymbol{\pi}} g + \kappa e^{-\frac{1}{2}\lambda^2 (T-t)} \begin{pmatrix} \mathbf{0} \\ -e^{-\widetilde{\pi}^{(n)}} \end{pmatrix} = \mathbf{0}.$$

The first $n-1$ components yield the equation

$$\overline{\boldsymbol{\Omega}}^{-1} (\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}}) + \begin{pmatrix} \widetilde{\pi}^{(1)} \\ \vdots \\ \widetilde{\pi}^{(n-1)} \end{pmatrix} + \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega} \widetilde{\pi}^{(n)} = 0$$

and hence

$$\begin{pmatrix} \widetilde{\pi}^{(1)} \\ \vdots \\ \widetilde{\pi}^{(n-1)} \end{pmatrix} = -\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}}) - \overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}\ \widetilde{\pi}^{(n)}. \tag{1.8}$$

Note that this relation also holds in the standard Merton problem. The difference is the equation that $\widetilde{\pi}^{(n),*}$ satisfies. The last line reads

$$(\mu_n - r)g + \boldsymbol{\omega}^T \begin{pmatrix} \widetilde{\pi}^{(1)} \\ \vdots \\ \widetilde{\pi}^{(n-1)} \end{pmatrix} g + \sigma_n^2 \widetilde{\pi}^{(n)} g - \kappa e^{-\frac{1}{2}\lambda^2(T-t)} \cdot e^{-\widetilde{\pi}^{(n)}} = 0.$$

Using the previous equation, we get the following equation for $\widetilde{\pi}^{(n)}$:

$$(\mu_n - r)g - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}})\ g + \left(\sigma_n^2 - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}\right)\ g\ \widetilde{\pi}^{(n)} - \kappa e^{-\frac{1}{2}\lambda^2(T-t)} \cdot e^{-\widetilde{\pi}^{(n)}} = 0. \tag{1.9}$$

To show that for $g > 0$ this equation has exactly one solution $\widetilde{\pi}^{(n),*}$, we need the following

Lemma 3.

$$\sigma_n^2 - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega} > 0$$

Proof. Recall that

$$\boldsymbol{\Omega} = \begin{pmatrix} \overline{\boldsymbol{\Omega}} & \boldsymbol{\omega} \\ \boldsymbol{\omega}^T & \sigma_n^2 \end{pmatrix},$$

and that we assume $\boldsymbol{\Omega}$ to be strictly positive definite. By the result from appendix 1.A, the bottom right entry of $\boldsymbol{\Omega}^{-1}$ is given by $(\sigma_n^2 - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega})^{-1}$. Since $\boldsymbol{\Omega}^{-1}$ is positive definite, this entry has to be positive, which proves the lemma. □

It follows that the left hand side of (1.9) is an increasing continuous function in $\widetilde{\pi}^{(n)}$ with limit ∞ as $\widetilde{\pi}^{(n)} \to \infty$ and with limit $-\infty$ as $\widetilde{\pi}^{(n)} \to -\infty$. Therefore we immediately get

Lemma 4. *For every $g > 0$, equation (1.9) has a unique solution $\pi^{(n),*}$.*

Substituting into (1.8) then yields the remaining components $\widetilde{\pi}^{(1),*}$, ..., $\widetilde{\pi}^{(n-1),*}$ of the optimal trading strategy.

It is possible to write $\widetilde{\pi}^{(n),*}$ in terms of the Lambert W-function, the inverse function of xe^x for $x \geq 0$. We let L (instead of W) denote the Lambert W-function for notational purposes. If we let

$$A \triangleq (\mu_n - r) - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}(\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}}), \quad B \triangleq \sigma_n^2 - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}, \quad x \triangleq \widetilde{\pi}^{(n),*}, \tag{1.10}$$

then (1.9) becomes $Ag + Bgx - \kappa e^{-\frac{1}{2}\lambda^2(T-t)} \cdot e^{-x} = 0$. This can be rewritten as $\left(x + \frac{A}{B}\right) e^{x + \frac{A}{B}} = \frac{\kappa e^{-\frac{1}{2}\lambda^2(T-t)}}{Bg} e^{\frac{A}{B}}$. The result is

$$\begin{aligned}
\widetilde{\pi}^{(n),*} =& L\left(\frac{\kappa e^{-\frac{1}{2}\lambda^2(T-t)}}{Bg} e^{\frac{A}{B}}\right) - \frac{A}{B} \\
=& L\left(\frac{\kappa e^{-\frac{1}{2}\lambda^2(T-t)}}{\left[\sigma_n^2 - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}\right] g} \exp\left\{\frac{(\mu_n - r) - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}(\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}})}{\sigma_n^2 - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}}\right\}\right) - \\
& - \frac{(\mu_n - r) - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}(\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}})}{\sigma_n^2 - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} \boldsymbol{\omega}}.
\end{aligned} \tag{1.11}$$

It is interesting to observe that there exists a value for κ such that $\pi_t^{(n),*} = 0$ for all $t \in [0, T]$. This value is given by

$$\kappa_0 = (\mu_n - r) - \boldsymbol{\omega}^T \overline{\boldsymbol{\Omega}}^{-1} (\overline{\boldsymbol{\mu}} - \overline{\boldsymbol{r}}),$$

and for this value of κ we have $g(t) = e^{-\frac{1}{2}\lambda^2(T-t)}$ and hence $U = V^{(n-1)}$. We can interpret this as the fact that when $\kappa = \kappa_0$, the additional opportunity of investing in $S^{(n)}$ is neutralized by the default risk and the associated loss of $\pi^{(n),*}$.

Furthermore, it is worth noting that κ_0 is the expected excess return rate (over r) under the minimal entropy martingale measure for $S^{(1)}$, ..., $S^{(n)}$, i.e. the measure Q which is equivalent to $\mathbb{P}$, under which $S^{(1)}$, ..., $S^{(n)}$ grow at rate r, and whose entropy with respect to $\mathbb{P}$ is minimized among all such measures.

Even though it is not possible to obtain the solution of the ODE (1.6), we can immediately determine an upper and a lower bound for g (and hence an upper bound for the

value function U). Obviously,

$$\begin{aligned}\inf_{\boldsymbol{\pi}\in\mathbb{R}^n} f(\boldsymbol{\pi}) &\geq g \inf_{\widetilde{\boldsymbol{\pi}}\in\mathbb{R}^n}\left\{\widetilde{\boldsymbol{\pi}}(\boldsymbol{\mu}-\boldsymbol{r})+\frac{1}{2}\widetilde{\boldsymbol{\pi}}^T\boldsymbol{\Omega}\widetilde{\boldsymbol{\pi}}\right\}\\ &= -g\cdot\frac{1}{2}(\boldsymbol{\mu}-r)^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu}-\boldsymbol{r}) = -\frac{1}{2}\Lambda^2\ g. \end{aligned} \tag{1.12}$$

Then a lower bound for g is given by the solution h_1 of the ODE

$$\begin{cases} h' - \left(\kappa+\frac{1}{2}\Lambda^2\right)\ h = 0,\\ h(T) = 1,\end{cases}$$

i.e. $h_1(t) = e^{-\left(\kappa+\frac{1}{2}\Lambda^2\right)(T-t)}$.

On the other hand,

$$\begin{aligned}\inf_{\boldsymbol{\pi}\in\mathbb{R}^n} f(\boldsymbol{\pi}) &\leq \inf_{(\pi^{(1)},\ \ldots,\ \pi^{(n-1)})\in\mathbb{R}^{n-1}} f(\pi^{(1)},\ \ldots,\ \pi^{(n-1)},\ 0)\\ &= -g\cdot\frac{1}{2}(\overline{\boldsymbol{\mu}}-\overline{\boldsymbol{r}})^T\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}}-\overline{\boldsymbol{r}}) + \kappa e^{-\frac{1}{2}\lambda^2(T-t)}\\ &= -\frac{1}{2}\lambda^2\ g + \kappa e^{-\frac{1}{2}\lambda^2(T-t)}. \end{aligned} \tag{1.13}$$

Therefore an upper bound for g is given by solution h_2 of the ODE

$$\begin{cases} h' - \left(\kappa+\frac{1}{2}\lambda^2\right)\ h = -\kappa e^{-\frac{1}{2}\lambda^2(T-t)}\\ h(T) = 1,\end{cases}$$

i.e. $h_2(t) = e^{-\frac{1}{2}\lambda^2(T-t)}$. This result is intuitively clear, since $V^{(n-1)}$, the value function of the Merton problem corresponding to an investment in $S^{(1)}, \ldots, S^{(n-1)}$ only, is a lower bound for the value function U.

We are now ready to prove

Theorem 5. *Equation (1.6) has a unique solution g on the interval $[0,T]$, and this solution satisfies*

$$e^{-\frac{1}{2}(\kappa+\Lambda^2)(T-t)} \leq g(t) \leq e^{-\frac{1}{2}\lambda^2(T-t)}$$

for all $t\in[0,T]$.

Proof. It is straightforward to see that the proof also holds for $t \in (-\infty, T]$, and with a very slight modification for $t \in \mathbb{R}$. Since the it is a variation of the standard existence and uniqueness theorem for ODEs, we only sketch the proof.

Consider the complete Banach space $\mathcal{B} = C([0,T],\ \|\cdot\|_\infty)$, and define the closed subset

$$\mathcal{S} \triangleq \left\{ g:\ [0,T] \to \mathbb{R} \ \middle|\ e^{-\frac{1}{2}(\kappa+\Lambda^2)(T-t)} \leq g(t) \leq e^{-\frac{1}{2}\lambda^2(T-t)} \right\}.$$

We define an operator $\mathbf{\Psi}:\ \mathcal{S} \to \mathcal{B}$ by

$$\mathbf{\Psi}(g)(t) \triangleq 1 + \int_t^T -\kappa g(s) + \inf_{\widetilde{\boldsymbol{\pi}} \in \mathbb{R}^n} \left\{ \widetilde{\boldsymbol{\pi}}^T(\boldsymbol{\mu} - \boldsymbol{r})\ g(s) + \frac{1}{2}\widetilde{\boldsymbol{\pi}}^T \mathbf{\Omega} \widetilde{\boldsymbol{\pi}}\ g(s) + \kappa e^{-\frac{1}{2}\lambda^2(T-s)} \cdot e^{-\widetilde{\pi}^{(n)}} \right\} ds. \tag{1.14}$$

By lemma 4, $\mathbf{\Psi}$ is well-defined. Using the inequalities (1.12) and (1.13) it is straightforward to check that for $g \in \mathcal{S}$, we have $e^{-\frac{1}{2}(\kappa+\Lambda^2)(T-t)} \leq \mathbf{\Psi}(g)(t) \leq e^{-\frac{1}{2}\lambda^2(T-t)}$, and hence $\mathbf{\Psi}$ maps $\mathcal{S}$ into itself.

We can write the integrand in equation (1.14) as

$$-\kappa g(s) + (\widetilde{\boldsymbol{\pi}}^*)^T(\boldsymbol{\mu} - \boldsymbol{r})\ g(s) + \frac{1}{2}(\widetilde{\boldsymbol{\pi}}^*)^T \mathbf{\Omega} \widetilde{\boldsymbol{\pi}}^*\ g(s) + \kappa e^{-\frac{1}{2}\lambda^2(T-s)} \cdot e^{-\widetilde{\pi}^{(n),*}}$$

with $\widetilde{\boldsymbol{\pi}}^*$ as in (1.11) and (1.8). It follows that the integrand satisfies a global Lipschitz condition with respect to g for all elements in $\mathcal{S}$. Starting with an arbitrary $g_0 \in \mathcal{S}$, we define a sequence of elements in $\mathcal{S}$ by

$$g_{n+1} \triangleq \mathbf{\Psi}(g_n), \quad n = 0,\ 1,\ 2,\ \ldots,$$

and a standard argument shows that this sequence converges uniformly to a function $g \in \mathcal{S}$ which is a solution of equation (1.6). □

Furthermore it is interesting to consider a slight variation of the previous problem. Suppose that upon default of $S^{(n)}$ the investor does not restrict himself to an investment in $S^{(1)}, \ldots, S^{(n-1)}$ and the money market account, but replaces $S^{(n)}$ by another defaultable stock with exactly the same dynamics as $S^{(n)}$. This could e.g. be the case when an

investor constantly wants to hold a certain number of mostly solid stocks in his portfolio, but in order to increase his expected return, he is willing to hold one defaultable stock at any given time.

For this new setup we can interpret the wealth process as a controlled jump diffusion (for details see e.g. Øksendal and Sulem (2007)). In this case, in analogy to (1.5) the HJB equation for U becomes

$$\begin{cases} \partial_t U + rw\ \partial_w U + \sup\limits_{\boldsymbol{\pi}\in\mathbb{R}^n} \left\{ \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ \partial_{ww}U + \boldsymbol{\pi}^T(\boldsymbol{\mu}-\boldsymbol{r})\ \partial_w U + \right. \\ \qquad\qquad\qquad\qquad \left. + \kappa\ \left[U(w-\pi^{(n)},\ t) - U\right] \right\} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad U(w,T) = u(w), \quad w\in\mathbb{R}. \end{cases}$$

The substitution $U(w,t) = u(we^{r(T-t)})\ g(t)$ then leads to the ODE for g

$$\begin{cases} g' - \kappa g + g\ \inf\limits_{\boldsymbol{\pi}\in\mathbb{R}^n} \left\{ \boldsymbol{\pi}^T(\boldsymbol{\mu}-\boldsymbol{r})a_t + \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ a_t^2 + \kappa e^{-a_t\pi^{(n)}} \right\} = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad g(T) = 1. \end{cases} \tag{1.15}$$

Note that the term to be minimized above, $\widetilde{\boldsymbol{\pi}}^T(\boldsymbol{\mu}-\boldsymbol{r}) + \frac{1}{2}\widetilde{\boldsymbol{\pi}}^T\boldsymbol{\Omega}\widetilde{\boldsymbol{\pi}} + \kappa e^{\widetilde{\pi}^{(n)}}$ depends on t, but not on g. As in the previous problem, we get the optimal strategy from the equations

$$(\mu_n - r) - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}}-\overline{\boldsymbol{r}}) + \left(\sigma_n^2 - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}\right)\ \widetilde{\pi}^{(n)} - \kappa e^{-\widetilde{\pi}^{(n)}} = 0,$$

$$\begin{pmatrix} \widetilde{\pi}^{(1)} \\ \vdots \\ \widetilde{\pi}^{(n-1)} \end{pmatrix} = -\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}}-\overline{\boldsymbol{r}}) - \overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}\ \widetilde{\pi}^{(n)}$$

Consequently,

$$\widetilde{\boldsymbol{\pi}}^{(n),*} = L\left(\frac{\kappa}{B}e^{\frac{A}{B}}\right) - \frac{A}{B}$$

with A, B are defined as in (1.10), and it follows that $\widetilde{\pi}^*$ is constant. If we let

$$m \triangleq \inf_{\widetilde{\boldsymbol{\pi}}\in\mathbb{R}} \left\{ \widetilde{\boldsymbol{\pi}}(\boldsymbol{\mu}-\boldsymbol{r}) + \frac{1}{2}\widetilde{\boldsymbol{\pi}}^T\boldsymbol{\Omega}\widetilde{\boldsymbol{\pi}} \right\} + \kappa e^{-\widetilde{\pi}^{(n)}},$$

we get the solution

$$g(t) = e^{(m-\kappa)(T-t)},$$

i.e. a similar result as in the standard Merton problem with $\frac{1}{2}\Lambda^2$ replaced by $\kappa - m$.

1.A Appendix: The Inverse of a Symmetric Matrix

The result from this appendix was used in section 1.2 and will be used several times in later sections. Given a symmetric matrix $\boldsymbol{X} \in \mathbb{R}^{(m+n)\times(m+n)}$ in block form,

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^T & \boldsymbol{C} \end{pmatrix},$$

with $\boldsymbol{A} \in \mathbb{R}^{m\times m}$, $\boldsymbol{B} \in \mathbb{R}^{m\times n}$, $\boldsymbol{C} \in \mathbb{R}^{n\times n}$, $\boldsymbol{A}$ and $\boldsymbol{C}$ symmetric and regular, we are interested in a representation for $\boldsymbol{X}^{-1}$.

We define the lower triangular matrix

$$\boldsymbol{L} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{O} \\ -\boldsymbol{B}^T\boldsymbol{A}^{-1} & \boldsymbol{I} \end{pmatrix}.$$

Then

$$\boldsymbol{L}\boldsymbol{X}\boldsymbol{L}^T = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{O}^T & -\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C} \end{pmatrix}. \tag{1.16}$$

Hence

$$\boldsymbol{L}^{-T}\boldsymbol{X}^{-1}\boldsymbol{L}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & \boldsymbol{O} \\ \boldsymbol{O}^T & \left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1} \end{pmatrix},$$

and therefore

$$\boldsymbol{X}^{-1} = \left(\begin{array}{c|c} \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}\left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1}\boldsymbol{B}^T\boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1} \\ \hline -\left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1}\boldsymbol{B}^T\boldsymbol{A}^{-1} & \left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1} \end{array}\right).$$

We note that the matrix multiplications in (1.16) correspond to using $\boldsymbol{A}$ to perform row operations on $\boldsymbol{X}$ to get zeros in the lower left block, and to perform the analogous column operations to get zeros in the upper right block.

However one can achieve the same result by using $\boldsymbol{C}$ instead of $\boldsymbol{A}$. Defining

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{I} & -\boldsymbol{C}^{-1}\boldsymbol{B} \\ \boldsymbol{O}^T & \boldsymbol{I} \end{pmatrix},$$

we get

$$\boldsymbol{U}\boldsymbol{X}\boldsymbol{U}^T = \begin{pmatrix} \boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^T & \boldsymbol{O} \\ \boldsymbol{O}^T & \boldsymbol{C} \end{pmatrix}.$$

Similarly as before we get a second representation for $\boldsymbol{X}^{-1}$, namely

$$\boldsymbol{X}^{-1} = \left(\begin{array}{c|c} \left(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^T\right)^{-1} & -\left(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^T\right)^{-1}\boldsymbol{B}\boldsymbol{C}^{-1} \\ \hline -\boldsymbol{C}^{-1}\boldsymbol{B}^T\left(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^T\right)^{-1} & \boldsymbol{C}^{-1} + \boldsymbol{C}^{-1}\boldsymbol{B}^T\left(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^T\right)^{-1}\boldsymbol{B}\boldsymbol{C}^{-1} \end{array}\right).$$

Obviously one can also combine these two representations. The most convenient one for future purpose is

$$\boldsymbol{X}^{-1} = \left(\begin{array}{c|c} \left(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^T\right)^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1} \\ \hline -\left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1}\boldsymbol{B}^T\boldsymbol{A}^{-1} & \left(-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B} + \boldsymbol{C}\right)^{-1} \end{array}\right).$$

Chapter 2

Indifference Pricing in a Reduced Form Model

In this chapter we use utility indifference pricing to price defaultable bonds and credit default swaps in a reduced form model. Default of the reference entity is triggered by the switching of a Poisson process N_t with $N_0 = 0$ and constant intensity κ. The assumption that κ is constant is mainly for notational purposes – we can easily generalize this setting for a time-dependent intensity κ_t. We assume that the investor can invest in a non-defaultable risky asset I (e.g. a stock index), the reference entity's stock S and the money market account. We restrict ourselves to only one default free risky asset, because as in sections 1.1 and 1.2 the results do not change significantly for a higher number of default free stocks.

The reference entity's default time is denoted by τ_d, formally

$$\tau_d \triangleq \inf\{t \geq 0 \mid N_t = 1\}.$$

I and S are assumed to have the dynamics

$$\begin{aligned} dI_t &= I_t \left(\mu_1 \, dt + \sigma_1 \, dB_t^{(1)} \right), \quad t \geq 0, \\ dS_t &= S_t \left(\mu_2 \, dt + \sigma_2 \, dB_t^{(2)} \right), \quad 0 \leq t < \tau_d, \\ S_t &= 0, \quad t \geq \tau_d, \end{aligned}$$

where $B_t^{(1)}$, $B_t^{(2)}$ are correlated Brownian motions with $dB_t^{(1)} dB_t^{(2)} = \rho dt$.

Throughout this chapter the recovery rate R of the defaultable bond is assumed to be a time invariant random variable independent of the driving Brownian motions.

We consider two different scenarios. In section 2.1 we examine the case when the investor does not invest in S at all, e.g. to avoid a possible loss due to default. This setup has the advantage that it is fairly mathematically tractable. We can obtain an explicit formula for bond prices and an implicit equation for CDS rates. In section 2.2 the investor is allowed to invest in both I and S. However, at time of default he loses the money invested in S at time τ_d. In this case, we have to solve the resulting optimization problems numerically, as the results in section 1.2 suggest.

2.1 Without Investment in the Defaultable Stock

2.1.1 The Defaultable Bond

The buyer of the bond receives a notional of F at maturity if the reference entity does not default before the maturity date T, or receives a percentage R of the notional at default if default occurs prior to maturity. For the seller of the bond the same rules apply, only that he has to make the above payments instead of receiving them.

Below we consider the case of the bond buyer. The case for the seller can be easily obtained by replacing F by $-F$. For later (e.g. equation (2.1)) we also let

$$\tilde{R}_t \triangleq -\frac{1}{\gamma F e^{r(T-t)}} \log \mathbb{E} \, e^{-\gamma R F e^{r(T-t)}},$$

so that we have $e^{-\gamma \tilde{R}_t F e^{r(T-t)}} = \mathbb{E}\ e^{-\gamma R F e^{r(T-t)}}$.

If we let $\tau \triangleq \tau_d \wedge T$, we model the dynamics of the wealth process by

$$d\overline{W}_t = \left[(\mu_1 - r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)} \quad \text{for } 0 < t < \tau \text{ and } t > \tau,$$

subject to $\overline{W}_\tau = \overline{W}_{\tau^-} + RF \cdot \mathbb{I}\{\tau_d \leq T\} + F \cdot \mathbb{I}\{\tau_d > T\}$.

Any trading strategy can be characterized by $\pi_t^{(1)}$, the amount of money invested in I at time t. The set of admissible trading strategies, again denoted by $\mathcal{A}$, is therefore defined as in definition 3 for $n = 1$. For convenience, in this section and section 2.1.2 we only write π instead of $\pi^{(1)}$.

The value function $\overline{V}$ corresponding to an investment in the defaultable bond is

$$\overline{V}(w, I, t) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[u(\overline{W}_T) \mid \overline{W}_t = w,\ I_t = I,\ t < \tau_d\right],$$

It is clear that $\overline{V}$ is independent of I, i.e. $\overline{V} = \overline{V}(w, t)$. The corresponding HJB equation is

$$\begin{cases} \partial_t \overline{V} + \sup\limits_{\pi \in \mathbb{R}} \left\{ [rw + (\mu_1 - r)\pi]\ \partial_w \overline{V} + \dfrac{1}{2}\sigma_1^2 \pi^2 \partial_{ww} \overline{V} \right\} + \\ \qquad\qquad\qquad\qquad + \kappa \left[V(w + \tilde{R}_t F,\ t) - \overline{V} \right] = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad \overline{V}(w, T) = u(w + F), \quad w \in \mathbb{R}\ . \end{cases} \tag{2.1}$$

Recall that V is the value function for the standard Merton investment problem (see equation (3.1)). The last term on the left hand side of (2.1) is due to a potential default and the corresponding switch to the standard Merton problem. In addition, applying the verification theorem from the appendix shows that any function satisfying equation (2.1) coincides with the value function $\overline{V}$.

Factoring out wealth (i.e. writing $\overline{V}(w, t) = u(w\ e^{r(T-t)})\ \overline{g}(t)$) yields the following equation for $\overline{g}$:

$$\begin{cases} \overline{g}' + \inf\limits_{\pi \in \mathbb{R}} \left\{ (\mu_1 - r)\pi a_t\ \overline{g} + \dfrac{1}{2}\pi^2 \sigma_1^2 a_t^2\ \overline{g} \right\} - \kappa\ \overline{g} + \kappa\ e^{-\frac{1}{2}\lambda^2 (T-t) + \tilde{R}_t F a_t} = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \overline{g}(T) = e^{-\gamma F}\ . \end{cases}$$

The infimum is attained at $\pi^{(1),*} = -\frac{\mu_1 - r}{2a_t\sigma_1^2} = -\frac{1}{2a_t}\lambda^2$ (with $\lambda = \frac{\mu_1 - r}{\sigma_1}$ as before), which leads to the linear ODE

$$\begin{cases} \overline{g}' - \left(\kappa + \dfrac{1}{2}\lambda^2\right)\ \overline{g} + \kappa\ e^{-\frac{1}{2}\lambda^2(T-t) + \tilde{R}_t F a_t} = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \overline{g}(T) = e^{-\gamma F}\ . \end{cases} \tag{2.2}$$

It is interesting to observe that even though the investor is exposed to a default risk, the optimal investment strategy does not change compared to the standard Merton investment problem. This is due to the fact that τ_d is independent of the stock price processes and that the expected change in utility at time τ_d is deterministic. This however will change in section 2.2, when the investor is also allowed to invest in the defaultable stock S.

Letting $\tilde{\kappa} = \kappa + \frac{1}{2}\lambda^2$ and $\tilde{\lambda}(t) = -\frac{1}{2}\lambda^2\ (T-t) + \tilde{R}_t F a_t$ for convenience, the solution differential equation can be written as

$$\begin{aligned} \overline{g}(t) &= e^{-\gamma F} \cdot e^{-\tilde{\kappa}(T-t)} + e^{\tilde{\kappa}t} \int_t^T \kappa e^{-\tilde{\kappa}s + \tilde{\lambda}(s)}\ ds \\ &= e^{-\gamma F} \cdot e^{-\tilde{\kappa}(T-t)} + \kappa e^{\kappa t} \cdot e^{-\frac{1}{2}\lambda^2(T-t)} \cdot \int_t^T e^{-\kappa s - \tilde{R}_s F \gamma e^{r(T-s)}}\ ds \\ &= e^{-\frac{1}{2}\lambda^2(T-t)} \cdot \left[e^{-\gamma F - \kappa(T-t)} + \kappa e^{\kappa t} \int_t^T e^{-\kappa s - \tilde{R}_s F \gamma e^{r(T-s)}}\ ds \right]\ . \end{aligned}$$

Interestingly, it is possible to rewrite this result in terms of an expectation over the default time as follows:

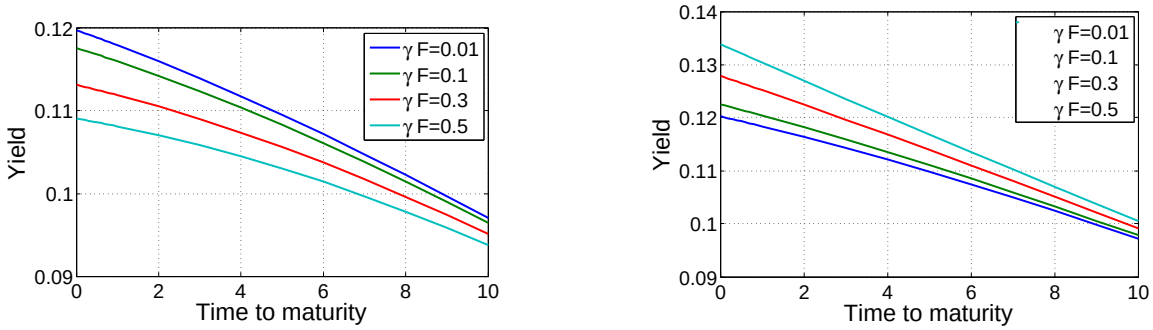
$$\overline{g}(t) = e^{-\frac{1}{2}\lambda^2(T-t)} \mathbb{E}\left[\exp\left\{ -\gamma \left(F \mathbb{I}_{\{\tau_d > T\}} + RFe^{r(T-\tau_d)}\ \mathbb{I}_{\{\tau_d \le T\}} \right) \right\} \middle| \tau_h < t < \tau_d \right]\ . \tag{2.3}$$

It is pleasing that an expectation over the risky bond's cash-flow accumulated to maturity arises in this context. This is of course a specific realization of the general duality result of Delbaen, Grandits, Rheinländer, Sampieri, Schweizer, and Stricker (2002). However, this duality result is not so simple to apply in the healthy region.

The indifference price p of the bond is given by the equation $\overline{V}(w-p,t) = V(w,t)$ yielding

$$p = -\frac{1}{a_t} \ln \frac{e^{-\frac{1}{2}\lambda^2(T-t)}}{\overline{g}(t)} = -\frac{1}{\gamma} e^{-r(T-t)} \ln \left(e^{-\gamma F} \cdot e^{-\kappa(T-t)} + \kappa \cdot \int_t^T e^{-\kappa(s-t) - \tilde{R}_s F \gamma e^{r(T-s)}}\ ds \right).$$

In Figure 2.1 we show the bond yield term structures with several levels of risk-aversion for both the seller and the buyer. Notice that as risk-aversion increases the buyer's yield increases as a more risk-averse investor demands a lower price and therefore a higher yield, while the opposite occurs for the seller. Interestingly, as the time to maturity grows, the spread decreases.



(a) Seller's Yields

(b) Buyer's Yields

Figure 2.1: The seller's and buyer's indifference yields for varying levels of risk-aversion in the distressed regime. The model parameters are: $r = 0.05$, $\mu_1 = 0.08$, $\sigma_1 = 0.2$, $\kappa = 0.1$. The values of μ_2, σ_2 and ρ are irrelevant.

2.1.2 The Credit Default Swap

Now suppose that the investor sells (or purchases) a CDS and receives (or pays) a continuous premium rate of A paid on a notional of F up until default time or maturity which ever occurs first. If default occurs first, the investor provides (or receives) a random payment of $(1-R)\,F$ (with $0 \leq R \leq 1$) and all future premium payments cease. Letting $\tau \triangleq \tau_d \wedge T$ as before, in this setup the wealth process has the dynamics

$$d\widetilde{W}_t = \begin{cases} \left[(\mu_1 - r)\pi_t^{(1)} + r\,\widetilde{W}_t + \epsilon AF\right]\,dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & 0 < t < \tau, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\,\widetilde{W}_t\right]\,dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & t > \tau, \end{cases}$$

subject to $\widetilde{W}_\tau = \widetilde{W}_{\tau^-} - \epsilon(1-R)F \cdot \mathbb{I}\{\tau_d \leq T\}$. Here $\epsilon = +1$ for the seller of the CDS and $\epsilon = -1$ for the buyer.

The value function $\tilde{V}$ corresponding to an investment in the CDS is defined as

$$\tilde{V}(w, I, t) \triangleq \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[u(\widetilde{W}_T) \;\middle|\; \widetilde{W}_t = w,\ I_t = I,\ \tau_h \leq t < \tau_d\right] .$$

Then considering that $\tilde{V}$ is independent of I, the HJB equation for the investor exposed to the CDS risk is

$$\begin{cases} \partial_t \tilde{V} + \sup_{\pi \in \mathbb{R}} \Big\{ [rw + \epsilon AF + (\mu_1 - r)\pi]\ \partial_w \tilde{V} + \frac{1}{2}\sigma_1^2 \pi^2 \partial_{ww} \tilde{V} + \\ \qquad\qquad + \kappa \left[V(w - \epsilon(1 - \tilde{R}_t)F,\ t) - \tilde{V} \right] \Big\} = 0 , \\ \qquad\qquad \tilde{V}(w, T) = u(w), \quad w \in \mathbb{R}, \end{cases} \tag{2.4}$$

where $\tilde{R}_t$ is defined as

$$\tilde{R}_t = -\frac{1}{\gamma \epsilon F e^{r(T-t)}} \log \mathbb{E}\ e^{-\gamma \epsilon R F e^{r(T-t)}}.$$

Letting $\tilde{V}(w, t) = u(we^{r(T-t)})\ \tilde{g}(t)$ leads to the following ODE for $\tilde{g}$:

$$\begin{cases} \partial_t \tilde{g} - (\kappa - \epsilon AF)\, a_t \tilde{g} + \inf_{\pi \in \mathbb{R}} \left\{ (\mu_1 - r) a_t \tilde{g} + \frac{1}{2}\pi^2 \sigma_1^2 a_t^2 \tilde{g} \right\} + \\ \qquad\qquad + \kappa e^{-\frac{1}{2}\lambda^2 (T-t) - \epsilon(1 - \tilde{R}_t) a_t} = 0 \\ \qquad\qquad \tilde{g}(T) = 1. \end{cases} \tag{2.5}$$

Again the infimum is attained at $\pi^* = -\frac{\mu_1 - r}{2 a_t \sigma_1^2} = -\frac{1}{2a_t}\lambda^2$, which leads to the equation

$$\begin{cases} \tilde{g}' - \left(\kappa + \frac{1}{2}\lambda^2 - \epsilon AF a_t\right)\ \tilde{g} + \kappa\ e^{-\frac{1}{2}\lambda^2 (T-t) - \epsilon(1 - \tilde{R}_t) F a_t} &= 0 \\ \overline{g}(T) &= 1 . \end{cases} \tag{2.6}$$

This ODE has the solution

$$\tilde{g}(t) = e^{-\frac{\lambda^2}{2}(T-t)} \left\{ e^{-\kappa(T-t)} \cdot e^{\epsilon AF \int_t^T a_u du} + \int_t^T e^{\epsilon F \left(A \int_t^s a_u du - (1 - \tilde{R}_s) a_s \right)}\ \kappa\ e^{-\kappa(s-t)}\ ds \right\} .$$

This can be simplified slightly by noticing that $\int_t^s a_u du = \frac{1}{r}(a_s - a_t)$; however, in its current form a natural interpretation arises akin to the result for the risky bond's value function in the distress region. In particular, it is easy to see that

$$\tilde{g}(t) = e^{-\frac{\lambda^2}{2}(T-t)} \mathbb{E}\left[\exp\left\{ -\gamma \left(\epsilon F A \int_t^{\tau_d \wedge T} e^{r(T-u)} du - \epsilon F (1 - R) e^{r(T - \tau_d)} \mathbb{I}_{\tau_d \leq T} \right) \right\} \;\middle|\; t < \tau_d \right] .$$

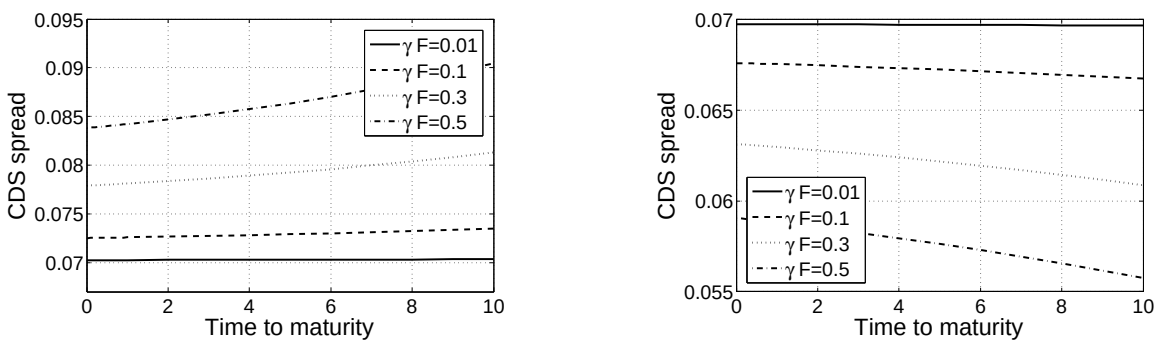
Once again this is a specific realization of the general duality results of Delbaen, Grandits, Rheinländer, Sampieri, Schweizer, and Stricker (2002).

Similar to the price of the defaultable bond, the indifference credit default swap spread is defined as the value $A = A(t)$ satisfying the equation $\tilde{V}(w,t) = V(w,t)$, leading to the equation

$$e^{\kappa t + \frac{\epsilon A F}{r} a_t} \int_t^T \kappa e^{-\kappa s - \frac{\epsilon A F}{r} a_s - \epsilon(1-\tilde{R}_s) F a_s} \, ds + e^{-\kappa(T-t) + \frac{\epsilon A F}{r}(a_t + \gamma)} = 1. \tag{2.7}$$

This time however, we cannot determine A analytically, but have to use numerical methods.

The plots in figure 2.2 show the seller's and buyer's CDS rates in the distressed regime using the same parameter values as in section 2.1.1, namely $r = 0.05$, $\mu_1 = 0.08$, $\sigma_1 = 0.2$, $\kappa = 0.1$, $R = 0.3$.



(a) Seller's CDS spreads

(b) Buyer's CDS spreads

Figure 2.2: The indifference CDS rate term structure for the buyer and seller in the distressed regime, determined from equation (2.7). See Figure 2.1 for the model parameters.

2.2 With Investment in the Defaultable Stock

In this section we consider the same setup as in section 2.1, except that the investor may now invest in I, S and the money market. Consequently, the set $\mathcal{A}$ of admissible trading strategies is defined as in definition 3 for $n = 2$. The method of solving the corresponding

optimization problems is similar as in the previous section, however we have to determine the solutions numerically. As in the previous section, we determine corporate bond prices and CDS spreads.

2.2.1 The Defaultable Bond

Letting $\tau \triangleq \tau_d \wedge T$, the dynamics of the wealth process are now given by

$$d\overline{W}_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi} + r\ \overline{W}_t\right]\ dt + \sigma_1 \pi_t^{(1)} dB_t^{(1)} + \sigma_2 \pi^{(2)} dB_t^{(2)}, & t < \tau, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)} \sigma_1 dB_t^{(1)}, & t > \tau, \end{cases}$$

subject to

$$\overline{W}_\tau = \overline{W}_{\tau^-} + \left(RF - \pi_t^{(2)}\right) \cdot \mathbb{I}\{\tau_d \leq T\} + F \cdot \mathbb{I}\{\tau_d > T\}.$$

The value function $\overline{U}$ corresponding to an investment in the defaultable bond is

$$\overline{U}(w, I, S, t) = \sup_{\pi \in \mathcal{A}} \mathbb{E}\left[u(\overline{W}_T) \mid \overline{W}_t = w,\ I_t = I,\ S_t = S,\ t < \tau_d\right].$$

Note that in contrast to the previous section, $\overline{U}$ formally depends on S as well. However, as before it can easily be seen that $\overline{W}_t$ does not depend on I and S, and therefore the corresponding HJB equation is

$$\begin{cases} \partial_t \overline{U} + \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \Big\{ \left[rw + (\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}\right]\ \partial_w \overline{U} + \frac{1}{2} \boldsymbol{\pi}^T \boldsymbol{\Omega} \boldsymbol{\pi}\ \partial_{ww} \overline{U} + \\ \qquad\qquad + \kappa \left[V(w - \pi^{(2)} + \tilde{R}_t F,\ t) - \overline{U}\right] \Big\} = 0, \\ \qquad\qquad\qquad\qquad \overline{U}(w, T) = u(w + F), \quad w \in \mathbb{R}\ . \end{cases} \tag{2.8}$$

The last term on the left hand side of this equation is due to a potential default and the corresponding switch between the state of no default to the state of default with the corresponding payment. In addition, applying the verification theorem from the appendix shows that any sufficiently integrable function satisfying equation (2.8) coincides with the value function $\overline{U}$.

Writing $\overline{V}(w,t) = u(w\ e^{r(T-t)})\ \overline{g}(t)$ this time yields the following equation for $\overline{g}$:

$$\begin{cases} \overline{g}' - \kappa\overline{g} + \inf\limits_{\boldsymbol{\pi}\in\mathbb{R}^2} \left\{ (\boldsymbol{\mu}-\boldsymbol{r})\boldsymbol{\pi}^T\ a_t\overline{g} + \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ a_t^2\overline{g} + \kappa\ e^{-\frac{1}{2}\lambda^2(T-t)+(\tilde{R}_tF-\pi^{(2)})a_t} \right\} = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overline{g}(T) = e^{-\gamma F}\ . \end{cases}$$

Similar to section 1.2, the optimal strategy is given implicitly by the equations

$$(\mu_n - r)\overline{g} - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}}-\overline{\boldsymbol{r}})\ \overline{g} + \left(\sigma_n^2 - \boldsymbol{\omega}^T\overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}\right)\ \widetilde{\pi}^{(n)}\ g - \kappa e^{-\widetilde{\pi}^{(n)}-a_t\tilde{R}F} = 0,$$

$$\begin{pmatrix} \widetilde{\pi}^{(1)} \\ \vdots \\ \widetilde{\pi}^{(n-1)} \end{pmatrix} = -\overline{\boldsymbol{\Omega}}^{-1}(\overline{\boldsymbol{\mu}}-\overline{\boldsymbol{r}}) - \overline{\boldsymbol{\Omega}}^{-1}\boldsymbol{\omega}\ \widetilde{\pi}^{(n)},$$

which for $n = 2$ becomes

$$(\mu_2 - r)\overline{g} - \rho\frac{\sigma_2}{\sigma_1}\ (\mu_1 - r)\overline{g} + (1-\rho^2)\sigma_2^2\widetilde{\pi}^{(2)}\overline{g} - \kappa e^{-\widetilde{\pi}^{(2)}-a_t\tilde{R}_tF} = 0$$

$$\widetilde{\pi}^{(1)} = -\frac{1}{\sigma_1^2}\ (\mu_1 - r) - \rho\ \frac{\sigma_2}{\sigma_1}\ \widetilde{\pi}^{(2)}.$$

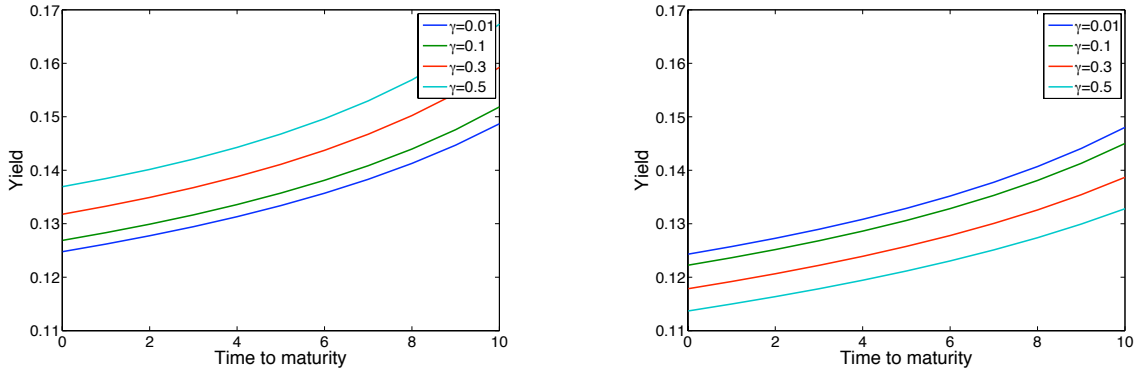
In contrast to section 2.1, one can see from the formula in the first line that the optimal strategy differs from the optimal strategies of both the standard Merton investment problem and the investment problem with a defaultable stock (as in section 1.2).

Finally, the indifference price of the defaultable bond is given by

$$p = -\frac{1}{a_t}\ln\frac{U(t)}{\overline{U}(t)} = -\frac{1}{a_t}\ln\frac{g(t)}{\overline{g}(t)}, \tag{2.9}$$

which we have to compute numerically in contrast to the case with no investment in the defaultable stock.

Figure 2.2.1 shows the yield curves for varying levels of risk aversion according to equation (2.9). The values of the functions g and $\tilde{g}$ were determined numerically using the Matlab ODE45 solver. It is surprising to see the significantly different shape of the yield curves compared to the ones in section 2.1.1. This difference obviously has to be due to the investment opportunity in the defaultable stock, and it would be interesting to get an intuitive understanding of this behaviour.



(a) Buyer's yields

(b) Seller's yields

Figure 2.3: The seller's and buyer's indifference yields for varying levels of risk-aversion in the distressed regime. The model parameters are: $r = 0.05$, $\mu_1 = 0.08$, $\mu_2 = 0.1$, $\sigma_1 = 0.2$, $\sigma_2 = 0.25$, $\rho = 0.5$, $\kappa = 0.1$, $R = 0.3$.

2.2.2 The Credit Default Swap

Letting $\tau \triangleq \tau_d \wedge T$ as before, the dynamics of the wealth process are now given by

$$d\widetilde{W}_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi} + r\ \widetilde{W}_t + \epsilon A F\right]\ dt + \pi_t^{(1)} \sigma_1\ dB_t^{(1)} + \pi^{(2)} \sigma_2\ dB_t^{(2)}, & t < \tau, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \widetilde{W}_t\right]\ dt + \pi_t^{(1)} \sigma_1\ dB_t^{(1)}, & t > \tau, \end{cases}$$

subject to

$$W_\tau = W_{\tau^-} - \left(\epsilon(1-R)F + \pi_t^{(2)}\right) \cdot \mathbb{I}\{\tau_d \leq T\}.$$

The value function corresponding to an investment in the credit default swap is

$$\tilde{U}(w,t) \triangleq \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}\left[u(\widetilde{W}_T) \mid \widetilde{W}_t = w,\ t < \tau_d\right].$$

Then the HJB equation for the investor exposed to the CDS risk is

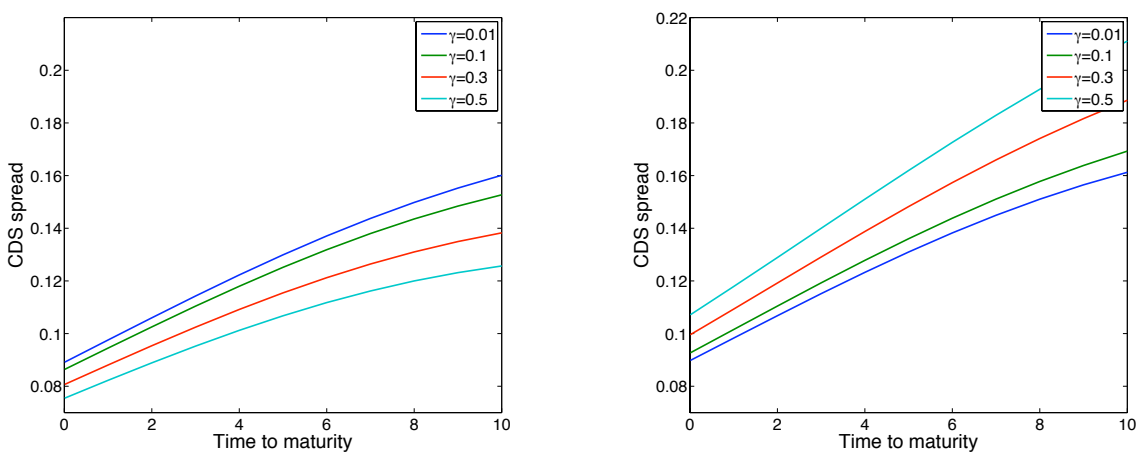
$$\begin{cases} \partial_t \tilde{U} + \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \Big\{ \left[rw + \epsilon A F + (\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}\right]\ \partial_w \tilde{U} + \frac{1}{2} \boldsymbol{\pi}^T \boldsymbol{\Omega} \boldsymbol{\pi}\ \partial_{ww} \tilde{U} + \\ \qquad\qquad + \kappa \left[V\left(w - \epsilon(1 - \tilde{R}_t)F - \pi^{(2)},\ t\right) - \tilde{V}\right] \Big\} = 0\ , \\ \qquad\qquad\qquad\qquad \tilde{U}(w,T) = u(w), \quad w \in \mathbb{R}\ . \end{cases} \tag{2.10}$$

The substitution $\tilde{V}(w,t) = u(we^{r(T-t)})\ \tilde{g}(t)$ leads to the following ODE for $\tilde{g}$:

$$\begin{cases} \partial_t \tilde{g} - (\kappa - \epsilon AF)\, a_t \tilde{g} + \inf\limits_{\boldsymbol{\pi}\in\mathbb{R}^2} \Big\{ (\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}\ a_t \tilde{g} + \dfrac{1}{2}\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi}\ a_t^2 \tilde{g} + \\ \qquad\qquad + \kappa e^{-\frac{1}{2}\lambda^2(T-t) - \left(\epsilon(1-\tilde{R}_t)F + \pi^{(2)}\right)a_t} \Big\} = 0 \\ \tilde{g}(T) = 1, \end{cases} \tag{2.11}$$

and the indifference CDS rate is again given by the implicit equation $\tilde{g}(t;A) = g(t)$. Here we have chosen the notation $\tilde{g}(t;A)$ for $\tilde{g}$ to emphasize its dependence on the parameter A.

The following plots show the buyer's and seller's CDS rates for the same parameters as in section 2.2.1.



(a) Buyer's CDS rates

(b) Seller's CDS rates

Figure 2.4: The seller's and buyer's CDS rates for varying levels of risk-aversion in the distressed regime. The model parameters are: $r = 0.05$, $\mu_1 = 0.08$, $\mu_2 = 0.1$, $\sigma_1 = 0.2$, $\sigma_2 = 0.25$, $\rho = 0.5$, $\kappa = 0.1$, $R = 0.3$.

2.A Appendix: Verification Theorems

In the sections 2.1 and 2.2 we determined the solutions of several optimization problems by formally solving the corresponding HJB equations. In this appendix we verify that these solutions indeed coincide with the value functions. We treat all cases simultaneously (i.e. bond, CDS, with and without investment in the defaultable stock).

We assume the same model as in sections 2.1 and 2.2. Let $\tau \triangleq \tau_d \wedge T$. Then for $\boldsymbol{\pi} \in \mathcal{A}$ we assume that the wealth process W_t has the following dynamics:

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ W_t + \epsilon A F\right]\ dt + \pi_t^{(1)} \sigma_1 dB_t^{(1)} + \pi_t^{(2)} \sigma_2 dB_t^{(2)}, & t < \tau, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ W_t\right]\ dt + \pi_t^{(1)} \sigma_1 dB_t^{(1)}, & t > \tau, \end{cases}$$

subject to

$$W_\tau \ = \ W_{\tau^-} + \left(-\pi^{(2)} + R_1\right) \cdot \mathbb{I}\{\tau < T\} + R_2 \cdot \mathbb{I}\{\tau = T\}.$$

A is a constant and corresponds to a continuous payment made ($\epsilon = -1$) or received ($\epsilon = +1$) up to time τ, or making/receiving no continuous payments at all ($\epsilon = 0$). R_1 is a time-independent random variable independent of $B_t^{(1)}$, $B_t^{(2)}$ and corresponds to a payment made/received at time τ_d, if $\tau_d < T$. Finally, R_2 is a deterministic constant and corresponds to a potential payoff at maturity T.

For $t \in [0, T]$ we let

$$\overline{U}(w, I, S, t) = \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}[u(W_T) \mid W_t = w,\ I_t = I,\ S_t = S,\ t < \tau_d].$$

If $\mathcal{A}$ is defined as in section 2.2, then $\overline{U}$ obviously is the same as the function $\overline{U}$ from section 2.2.1 for the case of the defaultable bond and coincides with $\tilde{U}$ from section 2.2.2 in the case of the credit default swap. If we restrict $\mathcal{A}$ to strategies $\boldsymbol{\pi}_t = (\pi_t^{(1)},\ \pi_t^{(2)})$ such that $\pi_t^{(2)} = 0$ for all $t \in [0, T]$, then $\overline{U}$ coincides with the value functions $\overline{U}$ from section 2.1.1 or $\tilde{U}$ from section 2.1.2.

Finally, if $\epsilon = R_1 = R_2 = 0$, then $\overline{U}$ from above coincides with the function U from section 1.2 for $n = 2$. Since the following theorem is almost identical for general n, it also

shows that the solution of the HJB equation in section 1.2 is indeed the value function from this section.

In all cases, $\overline{U}$ is independent of I and S, i.e. $\overline{U} = \overline{U}(w,t)$. We consider the corresponding HJB equation,

$$\begin{cases} \partial_t \overline{U} + \sup_{\boldsymbol{\pi}\in\mathbb{R}^2} \left\{ \mathcal{L}^{\boldsymbol{\pi}}\overline{U} + \kappa \left[V\left(w - \epsilon(1-\tilde{R}_t)F - \pi^{(2)},\ t\right) - \tilde{V} \right] \right\} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overline{U}(w,T) = u(w+R_2), \quad w \in \mathbb{R} \end{cases} \tag{2.12}$$

$$\tilde{R}_t = \frac{1}{\gamma\epsilon e^{r(T-t)}} \log \mathbb{E}\ e^{\gamma\epsilon R_1 e^{r(T-t)}},$$

and $\mathcal{L}^{\boldsymbol{\pi}}\overline{U}$ is given as

$$\mathcal{L}^{\boldsymbol{\pi}}\overline{U} = \left[rw + \epsilon AF + (\boldsymbol{\mu} - \boldsymbol{r})^T\boldsymbol{\pi}\right]\ \partial_w \overline{U} + \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ \partial_{ww}\overline{U}$$

Here V is the value function for the standard Merton investment problem with one risky asset. Moreover we let π_t^M be the corresponding optimal investment strategy.

Theorem 6. *Suppose there exists a function $H = H(w,t)$ which is a solution of (2.12) for $(w,t) \in \mathbb{R}\times[0,T]$ and which satisfies*

$$\mathbb{E}\int_0^T \boldsymbol{\pi}_t^2 \left(\partial_w H(W_t^{\boldsymbol{\pi}}, t)\right)^2\ dt < \infty$$

for all $\boldsymbol{\pi} \in \mathcal{A}$. Moreover, suppose that for all $(w,t) \in \mathbb{R}\times[0,T]$ there exists $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(w,t) \in \mathbb{R}$ such that

$$\mathcal{L}^{\boldsymbol{\pi}^*}H = \sup_{\boldsymbol{\pi}\in\mathbb{R}^n} \mathcal{L}^{\boldsymbol{\pi}}H.$$

Assume that the trading strategy $\overline{\boldsymbol{\pi}}_t$ defined by

$$\overline{\boldsymbol{\pi}}_t = \begin{cases} \boldsymbol{\pi}^*(W_{t^-}, t), & t \le \tau_d, \\ (\pi_t^M, 0), & t > \tau_d \end{cases}$$

is admissible. Then $\overline{U} = H$ for $(w,t) \in \mathbb{R}\times[0,T]$, and $\overline{\pi}$ is an optimal strategy, i.e. $\overline{U}(w,t) = \mathbb{E}_t\left[u\left(W_T^{\overline{\pi}}\right)\right]$.

Proof. Let H be as in the theorem, and let π be any admissible trading strategy. Instead of W_t consider the *continuous part* of the wealth process, i.e. the process $W_t^{(c)}$ satisfying

$$W_t^{(c)} = \begin{cases} W_t, & t < \tau, \\ W_t + \left(-\pi_\tau^{(2)} + R_1\right) \cdot \mathbb{I}\{\tau < T\} + R_2 \cdot \mathbb{I}\{\tau = T\}, & t \geq \tau. \end{cases}$$

Then $W_t^{(c)}$ is a diffusion and for $t \in [0, \tau)$ has the same dynamics as W_t. Hence Ito's lemma yields

$$H(W_\tau^{(c)},\ \tau) = H(w,t) + \int_t^\tau (\partial_t H + \mathcal{L}^{\boldsymbol{\pi}} H)\ ds + \int_t^\tau \partial_w H\ \left(\pi_s^{(1)}\sigma_1\ dB_s^{(1)} + \pi_s^{(2)}\sigma_2\ dB_s^{(2)}\right).$$

Since $\boldsymbol{\pi}$ is an arbitrary admissible strategy and noting that H solves the HJB equation, we always have $\partial_t H + \mathcal{L}^{\boldsymbol{\pi}} H \leq -\kappa \left[V(w + \tilde{R}_1 - \pi^{(2)},\ t) - H\right]$. Taking expectations on both sides makes the stochastic integral on the right hand side vanish and therefore yields

$$H(w,t) \geq \mathbb{E}_t\ H(W_\tau^{(c)}, \tau) + \mathbb{E}_t \int_t^\tau \kappa \left[V(W_s^{(c)} + \tilde{R}_1 - \pi^{(2)}, s) - H(W_s^{(c)}, s)\right]\ ds.$$

Here we use the notation $\mathbb{E}_t$ to abreviate the conditioning $W_t = w$.

Since the process $N_t - \kappa t$ is an $\mathcal{F}_t$-martingale, we have

$$\begin{aligned} &\mathbb{E}_t \int_t^\tau \kappa \left[V(W_s^{(c)} + \tilde{R}_1 - \pi^{(2)},\ s) - H(W_s^{(c)}, s)\right]\ ds \\ &\qquad = \mathbb{E}_t \int_t^\tau \left[V(W_s^{(c)} + \tilde{R}_1 - \pi^{(2)},\ s) - H(W_s^{(c)}, s)\right]\ dN_s \\ &\qquad\qquad = \mathbb{E}_t \left[\left(V(W_\tau^{(c)} + \tilde{R}_1 - \pi^{(2)},\ \tau) - H(W_\tau^{(c)}, \tau)\right) \cdot \mathbb{I}\{\tau_d \leq T\}\right], \end{aligned}$$

so we get

$$H(w,t) \geq \mathbb{E}_t \left[H(W_\tau^{(c)}, \tau) \cdot \mathbb{I}\{\tau_d > T\}\right] + \mathbb{E}_t \left[V(W_\tau^{(c)} + \tilde{R}_1, \tau) \cdot \mathbb{I}\{\tau_d \leq T\}\right]. \tag{2.13}$$

If $\tau_d > T$, then

$$H(W_\tau^{(c)}, \tau) = H(W_T^{(c)}, T) = u(W_T^{(c)} + R_2) = u(W_T),$$

and if $\tau_d \leq T$, then obviously $V(W_\tau^{(c)} + \tilde{R}_1 - \pi^{(2)}, \tau) = V(W_\tau, \tau) \geq \mathbb{E}_\tau u(W_T)$. Therefore, (2.13) implies $H(w,t) \geq \mathbb{E}_t\ [u(W_T)]$. Since this holds for any admissible strategy, it

follows that

$$H(w,t) \geq \overline{U}(w,t).$$

On the other hand, for $\pi = \overline{\pi}$ we get equality everywhere, and hence $H(w,t) = \overline{U}(w,t)$.

$\square$

Chapter 3

Indifference Pricing in a Structural Model of Default

3.1 Introduction

This section is an extension of the recent paper Leung, Sircar, and Zariphopoulou (2008), in which the authors introduce a new structural model to price corporate bonds. The reference entity's stock price and its firm value are modeled as correlated geometric Brownian motions, but in contrast to the common firm value models, the firm value is observable, but not tradable. Default of the firm is triggered by the asset value hitting a barrier D. The non-tradability of the firm's asset value makes the market incomplete. This contrasts with Sircar and Zariphopoulou (2007) where the authors analyze the effect of risk aversion within a reduced form approach.

We are interested in addressing how risk aversion and model uncertainty affect bond values and CDS rates. We adopt a similar setting to Leung, Sircar, and Zariphopoulou (2008) in the sense that we assume that the health of a company is measured by a credit-worthiness index (CWI; called the *firm's asset value* in Leung, Sircar, and Zariphopoulou (2008)). Since the health of a company is typically determined by more complex factors

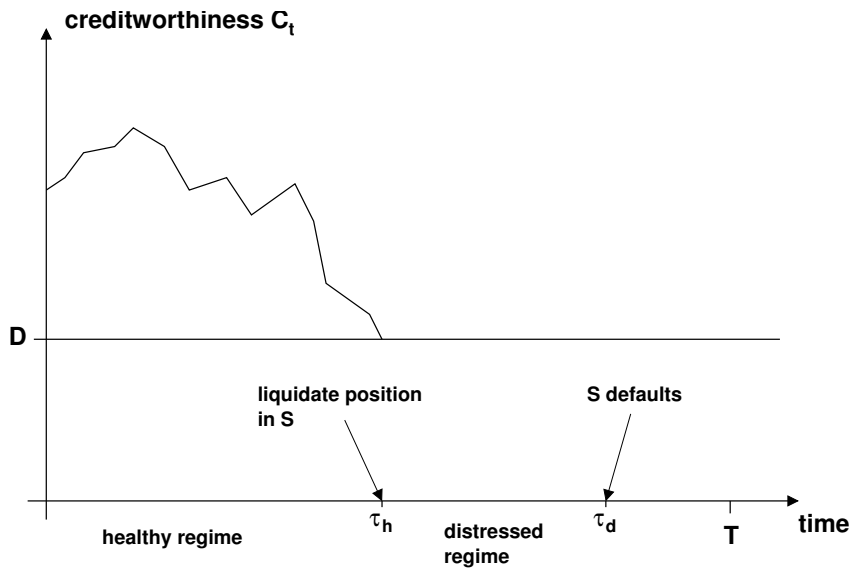
than the prices of its stocks and bonds, we assume that the CWI is not tradable. It is natural that the company's health will be correlated with its equity value, therefore we assume the CWI is positively correlated to the firm's stock price.

However we extend the model in Leung, Sircar, and Zariphopoulou (2008) in several aspects. Firstly, there is no reason to assume that the defaultable stock is the only available tradable asset. It is reasonable to assume that a real world investor is always able to invest in many liquid stocks, and more importantly, investors will try to diversify their portfolios. As a consequence, we consider a market in which the investor is additionally allowed to invest in a correlated non-defaultable index. Secondly, experience shows that it is not reasonable to assume that default of a company can be completely anticipated.

There are several ways to make the setup more realistic, one of which is to assume the default barrier to be time invariant, but unobservable. This model will be discussed in chapter 4. In this section, we consider an alternative setup: We assume that the barrier is a visible constant D, but after the CWI crosses it, the state of the company changes from *healthy* to *distressed*. At this point the company does not default yet, but enters a state of financial distress, in which default is triggered by an exogenous Poisson process. The model presented here is therefore not purely structural, but a hybrid model. Nevertheless, in this section we focus on indifference pricing when the firm's state is healthy, since the model for the distressed state is the same as in chapter 2.

In potential future work, it would be interesting to consider the case where in the distressed state the firm can either default or recover to the healthy state. This however would make the model less analytically tractable.

In this context, D can be interpreted as a rough upper estimate of an otherwise unknown default barrier, after whose hitting investors become nervous and withdraw their investments from the firm. Another interpretation of D is that of the level at which rating agencies downgrade the credit rating of the company.



3.2 The Model

Let a certain threshold $D > 0$ be given. Assuming that the CWI is above D at time 0, we first consider the state before the CWI hits D for the first time, which we shall call the *healthy regime* from now on. We work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, and let $\{(B_t^{(1)},\ B_t^{(2)},\ B_t^{(3)}) : 0 \leq t \leq T\}$ denote dependent $\mathbb{P}$-Wiener processes.

The non-defaultable index I, the reference entity's defaultable stock S and the creditworthiness index C are modeled as correlated geometric Brownian motions

$$\begin{aligned} dI_t &= I_t \left(\mu_1\ dt + \sigma_1\ dB_t^{(1)}\right) , \\ dS_t &= S_t \left(\mu_2\ dt + \sigma_2\ dB_t^{(2)}\right) , \\ dC_t &= C_t \left(\nu\ dt + \eta\ dB_t^{(3)}\right) \end{aligned}$$

with constant coefficients. For our purposes it is convenient to write the variance-covariance matrix of I, S, C in the form

$$\begin{pmatrix} \boldsymbol{\Omega} & \boldsymbol{\omega} \\ \boldsymbol{\omega}^T & \eta^2 \end{pmatrix} .$$

Here $\boldsymbol{\Omega}$ is the variance-covariance matrix of I, S, and $\boldsymbol{\omega} = (\rho_{13}\sigma_1\eta,\ \rho_{23}\sigma_2\eta)$, $dB_t^{(1)}dB_t^{(3)} = \rho_{13}\ dt$, $dB_t^{(2)}dB_t^{(3)} = \rho_{23}\ dt$.

Let

$$\tau_h \triangleq \inf\{t : \min_{0\le s\le t} C_t = D\}$$

be the first time that the CWI hits the threshold D. At this time S does not default yet. However the investor realizes that from now on, the firm is in a state of financial distress. As a consequence, he completely liquidates his investment in S and from thereon only invests in the money market and the non-defaultable index. Since S has not defaulted yet, it is resonable to assume that the investor can sell S at the current market price S_{τ_h}.

After C_t has hit D for the first time, the firm enters a state of financial distress, which will be called the *distressed regime* from now on. In the distressed regime default is triggered by the switching of a Poisson process N_t with $N_s = 0$ for $s \in [0, \tau_h]$ and hazard rate κ_t after τ_h. After time τ_h, N_t is independent of the Brownian motions $B_t^{(1)}$, $B_t^{(2)}$, $B_t^{(3)}$. We let τ_d denote the first arrival time of N_t after time τ_h, i.e.

$$\tau_d = \inf\{t > \tau_h \mid N_t = 1\}.$$

Since the investor has liquidated his position in the defaultable stock, the only sources of randomness in this state are $B_t^{(1)}$, and N_t, if invested in credit derivatives.

In this section, the natural filtration generated by the Wiener processes and the Poisson process is denoted $\mathcal{F} \triangleq \{\mathcal{F}_t : 0 \le t \le T\}$ where

$$\mathcal{F}_t = \sigma\left(\{(B_u^{(1)},\ B_u^{(2)},\ B_u^{(3)},\ N_u) : 0 \le u \le t\} \cup \mathcal{N}\right).$$

For the distressed regime we could alternatively choose the model from 2.2, with which we could avoid making the assumption that the investor liquidates his position in S at time τ_d. The changes however would be minimal.

3.3 The Investment Problem

We now define the set of admissible trading $\mathcal{A}$ strategies for our model. As previously, we require that an admissible strategy be $\mathcal{F}_t$-predictable. Note that in the healthy regime

this is the same as $\mathcal{F}_t$-adaptedness, but not in the distressed regime.

Definition 4. *An admissible trading strategy is an $\mathcal{F}_t$-predictable process $\boldsymbol{\pi}_t = (\pi_t^{(1)}, \pi_t^{(2)})$ satisfying the following:*

(i)

$$\pi_t^{(2)} = 0 \quad \textit{for} \quad t > \tau_h,$$

(ii)

$$\int_0^T \boldsymbol{\pi}_t^2 \, dt < \infty \quad \textit{almost surely,}$$

(iii)

$$\mathbb{E} \int_0^T \boldsymbol{\pi}_t^2 \left(e^{-\gamma e^{r(T-t)} W_t^{\boldsymbol{\pi}}} \right)^2 \, dt < \infty.$$

We begin by maximizing the investor's terminal expected utility of wealth in the two regimes. When the investor is not invested in any credit derivatives, the dynamics of the wealth process are given by

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r \, W_t \right] \, dt + \pi_t^{(1)} \sigma_1 \, dB_t^{(1)} + \pi_t^{(2)} \sigma_2 \, dB_t^{(2)}, & t < \tau_h, \\ \left[(\mu_1 - r) \pi_t^{(1)} + r \, W_t \right] \, dt + \pi_t^{(1)} \sigma_1 dB_t^{(1)}, & t > \tau_h, \end{cases}$$

subject to $W_{\tau_h} = W_{\tau_h^-}$. Note that in this setup, the investor is not exposed to any default risk.

We start with utility maximization the distressed regime. Since the investor is not exposed to any default risk, he is in the situation of the standard Merton investment problem with a money market account and the risky asset I, whose value function is

$$V(w,t) = -\frac{1}{\gamma} e^{a_t w - \frac{1}{2}\lambda^2 (T-t)}, \tag{3.1}$$

using the notation $a_t = -\gamma e^{r(T-t)}$ and $\lambda = \frac{\mu_1 - r}{\sigma_1}$.

Now we maximize expected terminal utility in the healthy regime through investment in the index I, the defaultable asset S and the money-market account.

We define the value function

$$U(w, I, S, C, t) = \sup_{\pi \in \mathcal{A}} \mathbb{E} \left[u(W_T) \mid W_t = w, \; I_t = I, \; S_t = S, \; C_t = C, \; t < \tau_h \right] .$$

Note that U is defined on the domain $\mathcal{D} \triangleq \mathbb{R} \times [0,\infty)^2 \times [D,\infty) \times [0,T]$.

A standard argument shows that assuming U to be sufficiently regular, we expect U to satisfy the partial differential equation

$$\begin{cases} \partial_t U + \sup\limits_{\boldsymbol{\pi}\in\mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} U = 0, \\ U(w,I,S,C,T) = u(w), \quad w \in \mathbb{R}, \;\; C > D, \\ U(w,I,S,D,t) = V(w,t), \quad w \in \mathbb{R}, \;\; t \in [0,T], \end{cases} \tag{3.2}$$

where $\mathcal{L}^{\boldsymbol{\pi}} U = \mathcal{K} U + \mathcal{K}^{\boldsymbol{\pi}} U$ and

$$\begin{aligned} \mathcal{K} U &\triangleq rw\, \partial_w U + \mu_1 I\, \partial_I U + \mu_2 S\, \partial_S U + \nu C\, \partial_C U + \frac{1}{2}\sigma_1^2 I^2\, \partial_{II} U + \frac{1}{2}\sigma_2^2 S^2\, \partial_{SS} U + \\ &\quad + \frac{1}{2}\eta^2 C^2\, \partial_{CC} U + \rho_{12}\sigma_1\sigma_2 IS\, \partial_{IS} U + \omega_1 IC\, \partial_{IC} U + \omega_2 SC\, \partial_{SC} U, \\ \mathcal{K}^{\boldsymbol{\pi}} U &\triangleq \frac{1}{2}\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi}\, \partial_{ww} U + \boldsymbol{\pi}^T \left[(\boldsymbol{\mu} - \boldsymbol{r}) \partial_w U + \boldsymbol{\Omega}(I\, \partial_{wI} U,\; S\, \partial_{wS} U)^T + \boldsymbol{\omega} C\, \partial_{wC} U \right] . \end{aligned}$$

The first boundary condition in (3.2) is the obvious terminal condition, and the second boundary condition is due to the firm's switching to the distressed regime at time τ_h.

It is straightforward to see that U is independent of I and S, i.e. $U(w,I,S,C,t) = U(w,C,t)$. Therefore the two terms above simplify to

$$\begin{aligned} \mathcal{K} U &= rw\, \partial_w U + \nu C\, \partial_C U + \frac{1}{2}\eta^2 C^2\, \partial_{CC} U, \\ \mathcal{K}^{\boldsymbol{\pi}} U &= \frac{1}{2}\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi}\, \partial_{ww} U + \boldsymbol{\pi}^T \left[(\boldsymbol{\mu} - \boldsymbol{r})\, \partial_w U + \boldsymbol{\omega} C\, \partial_{wC} U \right] . \end{aligned}$$

Furthermore, a standard verification theorem (see appendix 3.A) shows that the solution of (3.2) coincides with the value function U.

The first order condition for the optimal investment in the risky assets $\boldsymbol{\pi}$ is

$$\boldsymbol{\Omega}\, \partial_{ww} U\, \boldsymbol{\pi} = -(\boldsymbol{\mu} - \boldsymbol{r})\, \partial_w U - \boldsymbol{\omega} C\, \partial_{wC} U \ ,$$

which yields

$$\boldsymbol{\pi}^* = -\frac{1}{\partial_{ww} U}\, \boldsymbol{\Omega}^{-1} \left[(\boldsymbol{\mu} - \boldsymbol{r})\, \partial_w U + \boldsymbol{\omega} C\, \partial_{wC} U \right] .$$

Due to the exponential utility assumption, wealth can be removed from (3.2) by writing $U(w, C, t) = u\left(w\, e^{r(T-t)}\right)\ g(C, t)$, and we get

$$\begin{cases} \partial_t g + \nu C \partial_C g + \frac{1}{2}\eta^2 C^2 \partial_{CC} g - \\ -\frac{1}{2g}\ [(\boldsymbol{\mu} - \boldsymbol{r})g + \boldsymbol{\omega} C\ \partial_C g]\ \ \boldsymbol{\Omega}^{-1}\ [(\boldsymbol{\mu} - \boldsymbol{r})g + \boldsymbol{\omega} C\ \partial_C g] = 0 \\ g(D, t) = e^{-\frac{(\mu_1 - r)^2}{2\sigma_1^2}(T-t)}\ , \\ g(C, T) = 1\ . \end{cases}$$

Finally, very much like in Zariphopoulou (2001) and Henderson and Hobson (2002b) we make a substitution of the form

$$g(C, t) = G^{\beta}(\ln \frac{C}{D}, T - t)\, e^{-\frac{1}{2}\Lambda^2\,(T-t)},$$

where

$$\Lambda^2 = (\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\Omega}^{-1} (\boldsymbol{\mu} - \boldsymbol{r})$$

and β is chosen such that the resulting PDE for G becomes linear. The PDE for G is

$$-\partial_\tau G + \left(\nu - \frac{1}{2}\eta^2 - \boldsymbol{\omega}^T \boldsymbol{\Omega}^{-1}(\boldsymbol{\mu} - \boldsymbol{r})\right)\ \partial_x G + \frac{1}{2}\eta^2\ \partial_{xx} G + \frac{1}{2}\ \frac{(\partial_x G)^2}{G}\ \left[(\beta - 1)\ \eta^2 - \beta\ \boldsymbol{\omega}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\omega}\right] = 0.$$

The appropriate choice for β is

$$\beta = \frac{1}{1 - \frac{1}{\eta^2}(\boldsymbol{\omega}^T \boldsymbol{\Omega}^{-1} \boldsymbol{\omega})}\ ,$$

and the corresponding equation for $G(x, \tau)$ is

$$\begin{cases} -\partial_\tau G + \tilde{\nu}\ \partial_x G + \frac{1}{2}\eta^2\ \partial_{xx} G = 0 \\ G(0, \tau) = e^{\frac{1}{2\beta}(\Lambda^2 - \lambda^2)\tau}\ , \\ G(x, 0) = 1\ . \end{cases}$$

Here, $\widetilde{\nu} = \nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu} - \boldsymbol{r})^T\ \boldsymbol{\Omega}^{-1}\ \boldsymbol{\omega}$ is the drift of the CWI under the *minimal entropy martingale measure*[1], which is discussed in detail in Fritelli (2000). Due to the boundary

[1]Since the process C_t is continuous the MEMM measure is equivalent to the minimal martingale measure (see Schweizer (1999)).

condition along the barrier $C = D$, which is inherited from the subproblem of optimizing in the distressed regime, G is not simply the probability of remaining in the healthy regime under the MEMM. The PDE can be solved by using standard techniques (see appendix 3.B) to get

$$\begin{aligned} G(x,\tau) &= 1 - \frac{x}{\eta\sqrt{2\pi}} \int_0^\tau \frac{e^{-(x+\widetilde{\nu}u)^2/(2\eta^2 u)}}{u^{3/2}} \left[1 - e^{+\frac{1}{2\beta}(\Lambda^2-\lambda^2)(\tau-u)}\right] du \\ &= q_t(T;\widetilde{\nu}) + e^{(\widehat{\nu}-\widetilde{\nu})x/\eta^2 + \frac{1}{2\beta}(\Lambda^2-\lambda^2)\tau}\left(1 - q_t(T;\widehat{\nu})\right) , \end{aligned} \tag{3.3}$$

where

$$q_t(s;\theta) \triangleq \mathbb{Q}^\theta(\tau > s \mid X_t = x), \quad \widehat{\nu} = \widetilde{\nu} + \sqrt{\widetilde{\nu}^2 + \eta^2 \cdot \frac{1}{\beta}(\Lambda^2 - \lambda^2)}.$$

Here, as usual, $\Phi(y)$ denotes the standard normal cdf and $\mathbb{Q}^\theta$ is a measure induced the Radon-Nikodym derivative process

$$\left.\frac{d\mathbb{Q}^\theta}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \exp\left\{-\left(\tfrac{\nu-\theta}{\eta}\right)^2 - \left(\tfrac{\nu-\theta}{\eta}\right) B_t^{(3)}\right\} .$$

In terms of G the optimal trading strategy can be written as

$$\boldsymbol{\pi}^* = \frac{1}{\gamma e^{r(T-t)}} \boldsymbol{\Omega}^{-1}\left[(\boldsymbol{\mu} - \boldsymbol{r}) + \beta\boldsymbol{\omega}\,\frac{\partial_x G}{G}\right].$$

As in the optimization problems in previous chapters, $\boldsymbol{\pi}^*$ consists of the Merton-part given by the first term of the right hand side above, and the correction term implicitly given in terms of G and $\partial_x G$.

3.4 The Defaultable Bond

The investor receives a notional of F at maturity if the reference entity does not default before the maturity date T, or receives a percentage R (recovery) of the notional at default if default occurs prior to maturity. Consequently, if we let $\tau_1 \triangleq \tau_h \wedge T$, $\tau_2 \triangleq \tau_d \wedge T$, the

dynamics of the wealth process are given by

$$d\overline{W}_t = \begin{cases} \left[(\boldsymbol{\mu}-\boldsymbol{r})^T\boldsymbol{\pi}_t + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t<\tau_1, \\ \left[(\mu_1-r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_1<t<\tau_2, \\ \left[(\mu_1-r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & t>\tau_2, \end{cases}$$

subject to

$$\overline{W}_{\tau_1} = \overline{W}_{\tau_1^-} + F\cdot\mathbb{I}\{\tau_1=T\},$$
$$\overline{W}_{\tau_2} = \overline{W}_{\tau_2^-} + RF\cdot\mathbb{I}\{\tau_2<T\} + F\cdot\mathbb{I}\{\tau_2=T\}.$$

Mainly for notational purposes we assume that N_t has a constant hazard rate κ. However the computations can easily be generalized to non-constant, deterministic hazard rates.

The recovery rate is assumed to be random, but time-invariant and independent of the driving Brownian motions. We use the notation

$$\tilde{R}_t \triangleq -\frac{1}{\gamma F e^{r(T-t)}}\ \log\mathbb{E}\ e^{-\gamma RFe^{r(T-t)}},$$

i.e. $e^{-\gamma\tilde{R}_tFe^{r(T-t)}} = \mathbb{E}\ e^{-\gamma RFe^{r(T-t)}}$.

The derivation of the corresponding value function is similar as for the investment problem, however the expression for the value function in the distressed regime is not as simple as in the pure investment problem.

The value function is defined as

$$\overline{U}(w,I,S,C,t) \triangleq \sup_{\pi\in\mathcal{A}}\mathbb{E}\left[u(\overline{W}_T)\mid \overline{W}_t=w,\ I_t=I,\ S_t=S,\ C_t=C,\ t<\tau_h\right]\ .$$

Then we expect $\overline{U}$ to satisfy the HJB equation

$$\begin{cases} \partial_t\overline{U} + \sup_{\boldsymbol{\pi}\in\mathbb{R}^2}\mathcal{L}^{\boldsymbol{\pi}}\overline{U} = 0, \\ \overline{U}(w,I,S,C,T) = u(w+F), \quad w\in\mathbb{R},\ C>D, \\ \overline{U}(w,I,S,D,t) = \overline{V}(w,t), \quad w\in\mathbb{R}, \end{cases} \tag{3.4}$$

where

$$\overline{V}(w,t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \mathbb{E}[u(\overline{W}_T) \mid \overline{W}_t = w,\ \tau_h \le t < \tau_d]$$

is the value function in the distressed regime and therefore the same as the $\overline{V}$ from section 2.1.1. Compared to equation (3.2) only the boundary conditions are modified. As in the investment problem, the verification theorem from the appendix guarantees that any solution of (3.4) coincides with the value function $\overline{U}$.

Once again, it is clear that $\overline{U}$ is independent of I and S. Writing

$$\overline{U}(w,C,t) = u(w\,e^{r(T-t)})\ \overline{G}^{\beta}(\ln\frac{C}{D},\ T-t)\ e^{-\frac{1}{2\beta}\Lambda^2(T-t)}$$

as before, implies

$$\begin{cases} -\partial_\tau\overline{G} + \widetilde{\nu}\ \partial_x\overline{G} + \dfrac{1}{2}\eta^2\ \partial_{xx}\overline{G} = 0\ , \\ \qquad\qquad \overline{G}(0,\tau) = e^{\frac{1}{2\beta}\Lambda^2\tau}\ \overline{g}(T-\tau)^{1/\beta}\ , \\ \qquad\qquad \overline{G}(x,0) = e^{-\frac{\gamma F}{\beta}}\ , \end{cases} \tag{3.5}$$

whose solution is

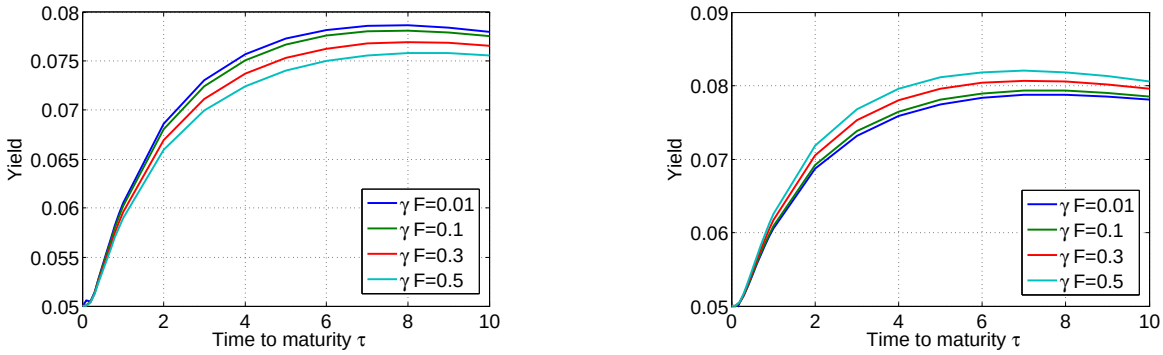
$$\overline{G}(x,\tau) = e^{-\frac{\gamma F}{\beta}} q_t(T;\widetilde{\nu}) + \frac{x}{\eta\ \sqrt{2\pi}} \int_0^\tau \frac{e^{-(x+\widetilde{\nu}u)^2/(2\eta^2 u)}}{u^{3/2}} e^{\frac{1}{2\beta}\ \Lambda^2(\tau-u)}\ \overline{g}(T-\tau+u)^{1/\beta}\ du. \tag{3.6}$$

Given (3.3) and (3.6), the indifference value $\overline{p}$ of the defaultable bond can be found by setting $U(w,C,t) = \overline{U}(w-\overline{p},C,t)$ from which we find

$$\overline{p}_t(T) = e^{-r\tau}\ \frac{\beta}{\gamma}\ \ln\frac{G(x,\tau)}{\overline{G}(x,\tau)}$$

with $x = \ln\frac{C}{D}$ and $\tau = T-t$. Unfortunately, we cannot simplify this expression any further; however, it is easy to numerically integrate using any standard quadrature routine.

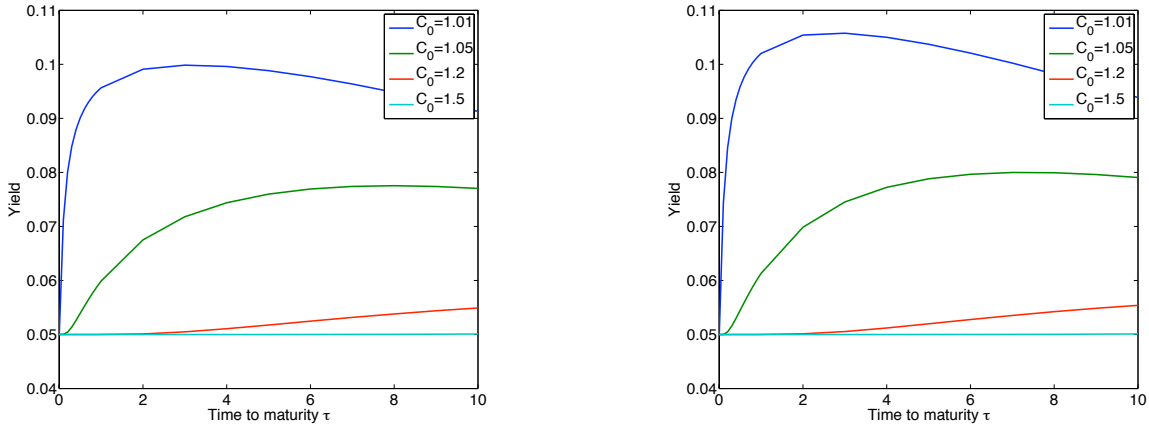
In Figure 3.1, the yield curves for different levels of risk-aversion are shown for a particular choice of parameters. Notice that in the healthy regime there is a definite hump shape in the risky yield despite the flat risk-free term structure. The hump is due to the non-zero recovery of 30% assumed in the example. Once again we observe the increasing/decreasing of the buyer's/seller's yields as risk-aversion increases.



(a) Seller's Yields (b) Buyer's Yields

Figure 3.1: The seller's and buyer's indifference yields for varying levels of risk-aversion in the healthy regime for $r = 0.05$, $\mu_1 = 0.08$, $\mu_2 = 0.1$, $\nu = 0.01$, $\sigma_1 = 0.2$, $\sigma_2 = 0.25$, $\eta = 0.05$ $\kappa = 0.1$, $\rho_{12} = 0.5$, $\rho_{13} = 0.3$, $\rho_{23} = 0.8$, $D = 1$, $C_0 = 1.05$, $R = 0.3$.

The plots in figure 3.2 show the yield curves for different levels of initial health C_0. Due to risk aversion, the seller will charge more for the bond than the buyer is willing to pay. Consequently, the buyer's yields are higher than those of the seller.



(a) Seller's yields (b) Buyer's yields

Figure 3.2: Buyer's and seller's yields for different values of initial health C_0. Risk aversion was chosen to be $\gamma F = 0.1$. The other parameters are the same as in figure 3.1.

3.5 The Credit Default Swap

We assume that the wealth process has the dynamics

$$d\widetilde{W}_t = \begin{cases} \left[(\boldsymbol{\mu}-\boldsymbol{r})^T\boldsymbol{\pi}_t + r\ \widetilde{W}_t + \epsilon AF\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t<\tau_1, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \widetilde{W}_t + \epsilon AF\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_1 < t < \tau_2, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \widetilde{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & t > \tau_2, \end{cases}$$

subject to $\widetilde{W}_{\tau_1} = \widetilde{W}_{\tau_1^-}$, $\widetilde{W}_{\tau_2} = \widetilde{W}_{\tau_2^-} - \epsilon(1-R)F\cdot\mathbb{I}\{\tau_2 < T\}$. Here $\epsilon = +1$ for the seller and $\epsilon = -1$ for the buyer of protection.

The value function in the healthy regime is

$$\tilde{U}(w,I,S,C,t) \triangleq \sup_{\pi\in\mathcal{A}} \mathbb{E}\left[u(\widetilde{W}_T)\ \middle|\ \widetilde{W}_t = w,\ I_t = I,\ S_t = S,\ C_t = C,\ t<\tau_h\right]\ ,$$

and again it is clear that $\tilde{U}$ is independent of I and S. The corresponding HJB equation is

$$\begin{cases} \partial_t\tilde{U} + \epsilon AF\partial_w\tilde{U} + \sup_{\boldsymbol{\pi}\in\mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}}\tilde{U} = 0, \\ \tilde{U}(w,D,t) = \tilde{V}(w,t), \\ \tilde{U}(w,C,T) = u(w), \quad C > D. \end{cases} \tag{3.7}$$

Here $\tilde{V}$ is the value function for the distressed regime, i.e.

$$\tilde{V}(w,t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \mathbb{E}[u(\widetilde{W}_T)\ |\ \overline{W}_t = w,\ \tau_h \le t < \tau_d].$$

As such, $\tilde{V}$ is the same as in section 2.1.2.

Equation (3.7) differs from equation (3.2) by the boundary condition along $C = D$ and the inclusion of the term $\epsilon AF\partial_w\tilde{U}$ representing the accumulation of premium payments. Assuming that $\tilde{U}$ has the form $\tilde{U}(w,C,t) = u\left(w\ e^{r(T-t)}\right)\ \tilde{G}^\beta\left(\ln\frac{C}{D}, T-t\right)\cdot e^{\psi(T-t)}$ with

$$\beta = \frac{1}{1-\frac{1}{\eta^2}(\boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega})}\ , \quad \psi(\tau) = -\frac{1}{2}\Lambda^2\tau - \epsilon\ \gamma\ \frac{AF}{r}\ e^{r\tau}$$

linearizes equation (3.7) resulting in

$$\begin{cases} -\partial_\tau \tilde{G} + \widetilde{\nu}\partial_x \tilde{G} + \dfrac{1}{2}\eta^2 \partial_{xx}\tilde{G} = 0 \ , \\ \qquad\qquad \tilde{G}(0,\tau) = e^{-\psi(\tau)/\beta} \cdot \tilde{g}(T-\tau)^{1/\beta} \ , \\ \qquad\qquad \tilde{G}(x,0) = 1 \ . \end{cases} \tag{3.8}$$

This can be solved as before to find

$$\tilde{G}(x,\tau) = q_t(T;\widetilde{\nu}) + \frac{x}{\eta\sqrt{2\pi}} \int_0^\tau \frac{e^{-(x+\widetilde{\nu}u)^2/(2\eta^2 u)}}{u^{3/2}} \, e^{-\psi(\tau-u)/\beta} \, \tilde{g}(T-\tau+u)^{1/\beta} \, du \ . \tag{3.9}$$

Armed with the solutions (3.3) and (3.9) the indifference CDS rate $A = A(C,t)$ makes the two value functions $U(w,C,t)$ and $\tilde{U}(w,C,t)$ equal and requires solving the non-linear equation

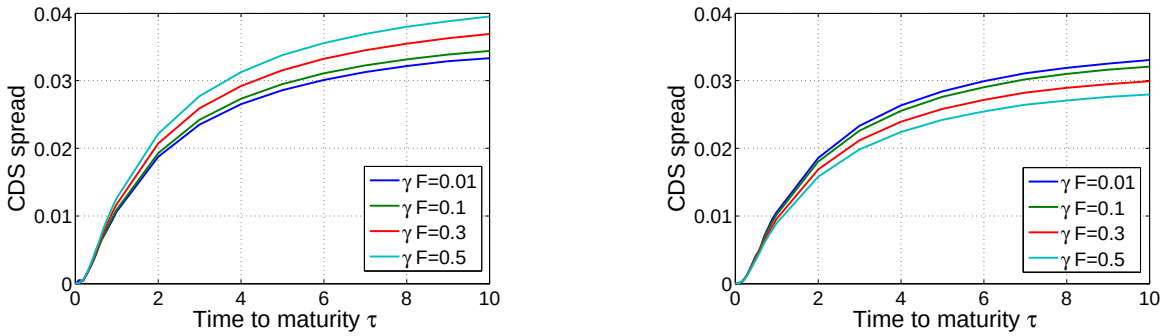
$$\epsilon\gamma \, F \, A = \frac{\beta \, r}{e^{r\tau}} \ln \frac{\tilde{G}(\ln \frac{C}{D},\tau;A)}{G(\ln \frac{C}{D},\tau)} \ . \tag{3.10}$$

The dependence of $\tilde{G}(x,\tau;A)$ on A is explicitly shown to emphasis the embedded non-linearity.

Figure 3.3 shows the seller's and buyer's CDS rates for the same parameters as in section 3.4. As risk-aversion increases, the seller's rates increase and the buyer's rates decrease. Generally, for the same level of risk aversion, the seller's rate is always higher than the buyer's rate. Unlike in the distressed regime, the spreads do indeed tend to zero for very short maturities; however, this occurs only at very short maturities. Once uncertainty in model parameters is accounted for, this steepening can be controlled not only by the proximity to the distress barrier, but also by the amount of model uncertainty. Chapter 5 addresses this issue.

The plots in figure 3.4 show the seller's and buyer's CDS spreads in the healthy regime for different levels of initial health C_0. The parameters are as before except that $\gamma = 0.2$ is fixed.

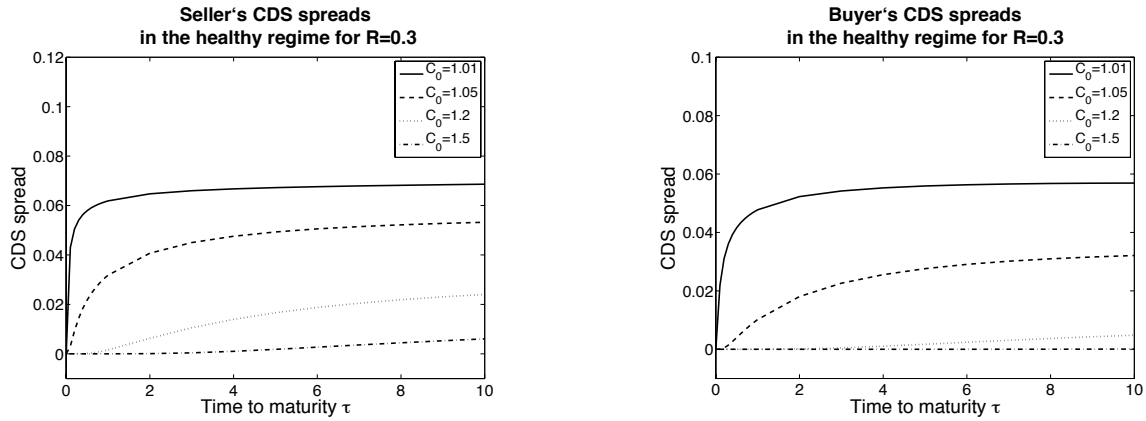
As expected, as the perceived health approaches the distress barrier, the CDS spread increases, while at every level of perceived health, the seller's rate is higher than the buyer's rate.



(a) Seller's CDS spreads

(b) Buyer's CDS spreads

Figure 3.3: The indifference CDS rate term structure for the buyer and seller in the healthy regime according to equation (3.10) for $r = 0.05$, $\mu_1 = 0.08$, $\mu_2 = 0.1$, $\nu = 0.01$, $\sigma_1 = 0.2$, $\sigma_2 = 0.25$, $\eta = 0.01$ $\kappa = 0.1$, $\rho_{12} = 0.5$, $\rho_{13} = 0.3$, $\rho_{23} = 0.8$, $C_0 = 1.05$, $D = 1$, $R = 0.3$



(a) Seller's CDS rates

(b) Buyer's CDS rates

Figure 3.4: CDS rates for $\gamma = 0.2$ and different levels of initial health C_0. The remaining parameters are $r = 0.05$, $\mu_1 = 0.08$, $\mu_2 = 0.1$, $\nu = 0.01$, $\sigma_1 = 0.2$, $\sigma_2 = 0.25$, $\eta = 0.01$ $\kappa = 0.1$, $\rho_{12} = 0.5$, $\rho_{13} = 0.3$, $\rho_{23} = 0.8$, $D = 1$, $R = 0.3$

3.A Appendix: Verification Theorem

In this section we prove that the solutions of the HJB equations in chapter 3 are indeed the value functions of the corresponding optimization problems. As in appendix 2.A, we

treat the cases of the investment problem, the defaultable bond and the credit default swap simultaneously.

Let $\tau_1 = \tau_h \wedge T$, $\tau_2 = \tau_d \wedge T$. We assume that the wealth process W_t has the dynamics

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ W_t + \epsilon AF\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t < \tau_1, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ W_t + \epsilon AF\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_1 < t < \tau_2, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ W_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & t > \tau_2, \end{cases}$$

subject to

$$\begin{aligned} W_{\tau_1} &= W_{\tau_1^-} \\ W_{\tau_2} &= W_{\tau_2^-} + R_1 \cdot \mathbb{I}\{\tau_2 < T\} + R_2 \cdot \mathbb{I}\{\tau_2 = T\}. \end{aligned}$$

A is a constant and corresponds to a continuous payment made ($\epsilon = -1$) or received ($\epsilon = +1$) up to time τ, or making/receiving no continuous payments at all ($\epsilon = 0$). R_1 is a time-independent random variable independent of the driving Brownian motions and corresponds to a payment made/received at time τ_d, if $\tau_d < T$. Finally, R_2 is a constant and corresponds to a potential payoff at maturity T.

For $t \in [0, T]$ we define

$$\overline{U}(w, I, S, C, t) = \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}[u(W_T) \mid W_t = w,\ I_t = I,\ S_t = S,\ t < \tau_h],$$

which in fact is independent of I and S. In the case of the defaultable bond, $\overline{U}$ is the same as in section 3.4, while for the investment problem and the CDS, $\overline{U}$ from above coincides with U from section 3.3 or $\tilde{U}$ from section 3.5. Moreover, for the theorem below, we let $\overline{\pi}$ denote the optimal trading strategy for the corresponding optimization problems from these sections.

We consider the HJB equation for $\overline{U}$,

$$\begin{cases} \partial_t \overline{U} + \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} \overline{U} = 0, \\ \qquad \overline{U}(w, C, T) = u(w + R_2), \quad w \in \mathbb{R},\ C > D, \\ \qquad \overline{U}(w, D, t) = \overline{V}(w, t), \quad w \in \mathbb{R},\ t \in [0, T], \end{cases} \tag{3.11}$$

with

$$\begin{aligned}\mathcal{L}^{\boldsymbol{\pi}}\overline{U} &= (rw+\epsilon AF)\ \partial_w\overline{U}+\nu C\ \partial_C\overline{U}+\frac{1}{2}\eta^2C^2\ \partial_{CC}\overline{U}+\\ &\quad+\frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ \partial_{ww}\overline{U}+\boldsymbol{\pi}^T\left[(\boldsymbol{\mu}-\boldsymbol{r})\ \partial_w\overline{U}+\boldsymbol{\omega}C\ \partial_{wC}\overline{U}\right]\ ,\end{aligned}$$

Theorem 7. *Suppose there exists a function $H=H(w,C,t)$ which solves (3.11) and which is sufficiently integrable in the sense that*

$$\mathbb{E}\int_0^T \boldsymbol{\pi}_t^2\,(\partial_w H)^2\ dt<\infty,\quad \mathbb{E}\int_0^{\tau_1}(\partial_C H)^2\ dt<\infty$$

for all $\boldsymbol{\pi}\in\mathcal{A}$. Suppose that for each $(w,C,t)\in\mathbb{R}\times(D,\infty)\times[0,T]$ there exists $\boldsymbol{\pi}^{**}=\boldsymbol{\pi}^{**}(w,C,t)\in\mathbb{R}^2$ such that

$$\mathcal{L}^{\boldsymbol{\pi}^{**}}H=\sup_{\boldsymbol{\pi}\in\mathbb{R}^2}\mathcal{L}^{\boldsymbol{\pi}}H. \tag{3.12}$$

Assume that the trading strategy defined by

$$\overline{\boldsymbol{\pi}}_t=\begin{cases}\boldsymbol{\pi}^{**}(W_t,C_t,t), & t<\tau_h,\\ (\overline{\pi}_t,0), & t\geq\tau_h,\end{cases}$$

is admissible. Then $\overline{U}=H$ for $(w,C,t)\in\mathbb{R}\times(D,\infty)\times[0,T]$, and $\overline{\boldsymbol{\pi}}$ is an optimal strategy, i.e. $\overline{U}(w,C,t)=\mathbb{E}_t\left[u\left(W_T^{\overline{\pi}}\right)\right]$.

Proof. Let $\overline{\boldsymbol{\pi}}$ and H as in the theorem, and let $\boldsymbol{\pi}\in\mathcal{A}$ be an arbitrary admissible strategy. For $t\in[0,\tau_1]$, $\overline{W}_t$ is a diffusion process. Writing τ instead of τ_1, we therefore get from Ito's lemma

$$\begin{aligned}H(W_\tau,C_\tau,\tau)&=H(w,C,t)+\int_t^\tau(\partial_tH+\mathcal{L}^{\boldsymbol{\pi}}H)\ ds+\\ &\quad+\int_t^\tau\partial_wH\cdot\left(\pi_s^{(1)}\sigma_1\ dB_s^{(1)}+\pi_s^{(2)}\sigma_2\ dB_s^{(2)}\right)+\partial_CH\ \eta C\ dB_s^{(3)}.\end{aligned}$$

Since $\boldsymbol{\pi}\in\mathcal{A}$ is an arbitrary strategy, we always have $\partial_tH+\mathcal{L}^{\boldsymbol{\pi}}H\leq 0$, so that taking expectations on both sides yields $H(w,C,t)\geq\mathbb{E}_t\ H(W_\tau,C_\tau,\tau)$. Making use of the fact that H is a solution of (3.11), we get

$$\begin{aligned}H(W_\tau,C_\tau,\tau)&=H(W_T,C_T,T)\cdot\mathbb{I}\{\tau_h>T\}+H(W_\tau,C_\tau,\tau)\cdot\mathbb{I}\{\tau_h\leq T\}\\ &=u(W_T)\cdot\mathbb{I}\{\tau_h>T\}+H(W_\tau,D,\tau)\cdot\mathbb{I}\{\tau_h\leq T\}\\ &=u(W_T)\cdot\mathbb{I}\{\tau_h>T\}+\overline{V}(W_\tau,\tau)\cdot\mathbb{I}\{\tau_h\leq T\}.\end{aligned}$$

Taking expectations on both sides and using the definition of $\overline{V}$ leads to

$$\begin{aligned}\mathbb{E}_t\ H(W_\tau, C_\tau, \tau) &\geq \mathbb{E}_t\left[u(W_T)\cdot \mathbb{I}\{\tau_h > T\}\right] + \mathbb{E}_t\left[u(W_T)\cdot \mathbb{I}\{\tau_h \leq T\}\right] \\ &= \mathbb{E}_t u(W_T),\end{aligned}$$

and hence $H(w, C, t) \geq \mathbb{E}_t u(W_T)$. Since $\boldsymbol{\pi}$ is an arbitrary admissible strategy, this implies $H(w, C, t) \geq \overline{U}(w, C, t)$. Now let $\boldsymbol{\pi} = \overline{\boldsymbol{\pi}}$. By the same argument we get equality in all the steps above, and therefore $H = \overline{U}$.

□

We still have to show that the value functions U, $\overline{U}$ and $\tilde{U}$ and the corresponding optimal trading strategies satisfy the conditions of theorem 7. Recall that we found that

$$U(w, C, t) = u(we^{r(T-t)}) \cdot e^{-\frac{1}{2}\Lambda^2(T-t)} \cdot G^\beta\left(\ln\frac{C}{D}, T-t\right) \tag{3.13}$$

$$\boldsymbol{\pi}^* = \frac{1}{\gamma e^{r(T-t)}}\boldsymbol{\Omega}^{-1}\left[(\boldsymbol{\mu} - \boldsymbol{r}) + \beta\boldsymbol{\omega}\ \frac{\partial_x G}{G}\right]. \tag{3.14}$$

For $\overline{U}$ and $\tilde{U}$ we have to replace G by $\overline{G}$ and $\tilde{G}$. From hereon, we focus on the proof for U, but the proofs for $\overline{U}$ and $\tilde{U}$ are analogous. We have

$$\begin{aligned}\partial_w U &= a_t\ u(we^{r(T-t)})\ e^{-\frac{1}{2}\Lambda^2(T-t)}\ G^\beta\left(\ln\frac{C}{D}, T-t\right) \\ C\ \partial_C U &= u(we^{r(T-t)})\ e^{-\frac{1}{2}\Lambda^2(T-t)}\ \beta G^{\beta-1}\cdot D\cdot \partial_x G.\end{aligned}$$

It is therefore straightforward to check that U satisfies the conditions from the verification theorem, if both G and $\partial_x G$ are bounded, and if additionally G is bounded from below by a positive constant. In particular this makes sure that $\boldsymbol{\pi}^*$ is bounded.

Recall that G is the solution of the equation

$$\begin{aligned}-\partial_t G + \tilde{\nu}\partial_x G + \frac{1}{2}\eta^2\partial_{xx}G &= 0, \\ G(x, 0) &= 1, \\ G(0, t) &= h(t),\end{aligned}$$

where h is continuously differentiable on $[0,T]$, bounded from below by a positive constant and $\lim_{t\to 0} h(t) = 1$. The claim for G then follows from the maximum and minimum principle of the heat equation. Furthermore, from the representation of G from appendix 3.B, it follows that $\partial_x G$ exists everywhere and is bounded along the boundary. Since $\partial_x G$ satisfies a heat equation as well, we get from the maximum and minimum principle, that $\partial_x G$ is bounded on $[0,\infty)\times[0,T]$.

3.B Appendix: The Heat Equation on the Quarter Plane

We would like to find a solution u to the heat equation

$$\begin{cases} \partial_t u + \nu\ \partial_x u + \frac{1}{2}\eta^2\ \partial_{xx} u = 0, \\ u(0,t) = g(t), \\ u(x,T) = f(x) \end{cases} \tag{3.15}$$

for $x \geq 0$ and $t \leq T$. We assume that f and g are continuous. Alternatively we can solve the equation

$$\begin{cases} -\partial_t \overline{u} + \nu\ \partial_x \overline{u} + \frac{1}{2}\eta^2\ \partial_{xx} \overline{u} = 0, \\ \overline{u}(0,t) = g(T-t), \\ \overline{u}(x,0) = f(x) \end{cases} \tag{3.16}$$

for $x \geq 0$, $t \geq 0$ and then let $u(x,t) = \overline{u}(x, T-t)$.

Assume that u is a solution of (3.15) and fix x and t. As introduced in section 3.3, for $\theta \in \mathbb{R}$ let $\mathbb{Q}^\theta$ be a measure under which a certain stochastic process has the dynamics $X_s \triangleq x + \theta(s-t) + \eta B^\theta_{s-t}$ where B^θ_s is a standard Brownian motion under $\mathbb{Q}^\theta$. Furthermore let $\tau \triangleq \inf\{s \geq t :\ X_s = 0\} \wedge T$.

Working under the measure $\mathbb{Q}^\nu$, we get from Ito's lemma

$$u(X_\tau, \tau) = u(x,t) + \int_t^\tau \left(\partial_t u + \nu\ \partial_x u + \frac{1}{2}\eta^2\ \partial_{xx} u\right)\ dt + \int_t^\tau \eta\ \partial_x u\ dB^\nu_s.$$

Taking expectations on both sides and using the fact that u solves the given heat equation yields

$$u(x,t) = \mathbb{E}^{\mathbb{Q}^\nu}\left[u(X_\tau,\tau)\right] = \mathbb{E}^{\mathbb{Q}^\nu}\left[g(\tau)\cdot\mathbb{I}\{\tau\le T\} + f(X_T)\cdot\mathbb{I}\{\tau>T\}\right]. \tag{3.17}$$

If f is a constant K as in this paper, then $\mathbb{E}^{\mathbb{Q}^\nu}[f(X_T)\cdot\mathbb{I}\{\tau>T\}]$ obviously simplifies to $K\cdot q_t(T;\nu)$, where

$$q_t(s;\theta) \triangleq \mathbb{Q}^\theta(\tau>s)$$

Under certain circumstances we can also simplify the first term on the right hand side of equation (3.17). Switching to the measure $\mathbb{Q}^0$ under which X_s has the dynamics $X_s = x+\eta B^0_{s-t}$ and applying the reflection principle, we get the well-known result that

$$q_t(s;\theta) = \Phi\left(\frac{x+\theta(s-t)}{\eta\sqrt{s-t}}\right) - e^{-2\theta x/\eta^2}\Phi\left(\frac{-x+\theta(s-t)}{\eta\sqrt{s-t}}\right).$$

We can compute the corresponding density $d_t(s;\theta)$ of τ by differentiating $1-q_t(s;\theta)$ with respect to s to get

$$\begin{aligned}
d_t(s;\theta) &= \frac{-1}{\sqrt{2\pi}}\left[\exp\left\{-\frac{1}{2}\left(\frac{x+\theta(s-t)}{\eta\sqrt{s-t}}\right)^2\right\}\cdot\left(-\frac{x}{2\eta(s-t)^{3/2}}+\frac{\theta}{2\eta\sqrt{s-t}}\right)\right.\\
&\quad\left.-e^{-2\theta x/\eta^2}\cdot\exp\left\{-\frac{1}{2}\left(\frac{-x+\theta(s-t)}{\eta\sqrt{s-t}}\right)^2\right\}\cdot\left(\frac{x}{2\eta(s-t)^{3/2}}+\frac{\theta}{2\eta\sqrt{s-t}}\right)\right]\\
&= \frac{1}{\sqrt{2\pi}}\frac{x}{\eta(s-t)^{3/2}}\cdot\exp\left(-\frac{[x+\theta(s-t)]^2}{2\eta^2(s-t)}\right)
\end{aligned}$$

We then have

$$\mathbb{E}^{\mathbb{Q}^\nu}\left[g(\tau)\cdot\mathbb{I}\{\tau\le T\}\right] = \int_t^T g(s)\ d_t(s;\nu)\ ds.$$

If the boundary condition at $x=0$ is of the form $u(0,t)=e^{L(T-t)}$ for some constant L, it follows from equation (3.18) that

$$\begin{aligned}
\mathbb{E}\left[g(\tau)\cdot\mathbb{I}\{\tau\le T\}\right] &= \frac{1}{\sqrt{2\pi}}\ e^{L(T-t)}\cdot\int_t^T\frac{x}{\eta(s-t)^{3/2}}\ e^{-L(s-t)}\cdot\exp\left\{-\frac{(x+\theta(s-t))^2}{2\eta^2(s-t)}\right\}\ ds\\
&= e^{L(T-t)}\cdot e^{\frac{1}{\eta^2}(x\sqrt{\theta^2+2\eta^2L}-x\theta)}\cdot\\
&\quad\cdot\frac{1}{\sqrt{2\pi}}\int_t^T\frac{x}{\eta(s-t)^{3/2}}\ \exp\left\{-\frac{(x+\sqrt{\theta^2+2\eta^2L}\ (s-t))^2}{2\eta^2(s-t)}\right\}\ ds\\
&= e^{L(T-t)+\frac{(\widehat{\theta}-\theta)x}{\eta^2}}(1-q_t(T;\widehat{\theta}))
\end{aligned}$$

with $\widehat{\theta} = \sqrt{\theta^2 + 2\eta^2 L}$.

Chapter 4

Indifference Pricing with Invisible Default Boundary

In this section we consider a similar structural model as in the previous section. The investor can invest in n tradable assets $S^{(1)}, \ldots, S^{(n)}$, which are modeled as correlated geometric Brownian motions with constant coefficients. The assets $S^{(1)}, \ldots, S^{(n-1)}$ are default free, while $S^{(n)}$ is defaultable. As before, the creditworthiness of $S^{(n)}$ is given by a strongly correlated creditworthiness index C_t, which we assume as observable, but non-tradable. Upon C_t hitting a critical threshold D for the first time, which we will denote by τ_h as before, the investor liquidates his position in $S^{(n)}$. The difference to the previous section is that here we assume the critical barrier D to be unobservable to the investor. We do however assume that the event of C hitting D is observable. This is e.g. consistent with interpreting D as a default barrier, since a default can normally be observed when it happens.

We investigate the following two scenarios:

- We interpret D as a default barrier. As a consequence, at time τ_h the investor makes/receives potential payments due to investments in credit derivatives. However, at τ_h the value of $S^{(n)}$ also drops to zero. Since this drop cannot be anticipated

by the investor due to D being unobservable, the investor loses the money he has invested in $S^{(n)}$, i.e. his wealth instantly drops by this amount. This case will be subsequently be called the *default case*.

- $S^{(n)}$ does not default yet at time τ_h, but the firm enters a state of financial distress in which default is triggered by a Poisson process. In this case there is no instant drop in the stock price, however the investor liquidates his position in $S^{(n)}$. After time τ_h, default is triggered by the switching of a Poisson process independent of the Brownian motions driving the stocks. This case is the analogue to the scenario investigated in chapter 3 and will be referred to as the *non-default case*. The model for the time interval $[\tau_h, T]$ is the same as for the distressed regime in chapter 3 and will also be referred to as such. The default time will again be denoted by τ_d.

The main mathematical difference between the two cases is that in the default case the change in wealth at time τ_h is not deterministic, but depends on the trading strategy. This additional term containing the control seems to make the corresponding optimization problem more difficult, which however is not the case, as will be shown. Nevertheless for the above reason, we first examine the non-default case, then the default case.

In a recent paper, Giesecke (2006) considers a similar setup in a structural framework and uses it to determine corporate bond prices and credit spreads. However, this paper and our work differ in many ways. Firstly, Giesecke works with the classical firm value model, in which the firm value is assumed to be an observable and tradable asset. The company defaults, when the firm value crosses the default barrier D for the first time. Secondly, he uses risk-neutral pricing. This setup has the advantage that many explicit results can be obtained. Even though in the context of indifference pricing, we are only able to obtain very limited explicit results, some of Giesecke's concepts are very useful.

When using this setup in the context of indifference pricing, as usual we have to solve two portfolio optimization problems and then find the indifference price that makes the two value functions equal. As we will see, the HJB equations corresponding to the value

functions can be reduced to heat equations in a natural way, very similar as in the case of the visible critical barrier. This time however, the boundary conditions of these equations are not of the Dirichlet type, but of a non-linear Robin type. This makes it impossible to obtain solutions in closed form. Despite this lack of explicit results for pricing purposes, the underlying portfolio optimization problems are interesting by themselves.

4.1 The Model

As in the previous sections, we model the dynamics of $S^{(1)}, \ldots, S^{(n)}, C$ as geometric Brownian motions, i.e.

$$\begin{aligned} dS_t^{(i)} &= S^{(i)} \left(\mu_1 \, dt + \sigma_1 \, dB_t^{(i)}\right), && i = 1, \ldots, n, \quad t \in [0, T] \\ dC_t &= C_t \left(\nu \, dt + \eta \, dB_t^{(n+1)}\right), && t \leq \tau_h, \end{aligned}$$

where $B_t^{(1)}, \ldots, B_t^{(n+1)}$ are correlated geometric Brownian motions. The covariance matrix of $S^{(1)}, \ldots, S^{(n)}, C$ will be denoted by

$$\begin{pmatrix} \boldsymbol{\Omega} & \boldsymbol{\omega} \\ \boldsymbol{\omega}^T & \eta^2 \end{pmatrix}.$$

We assume this matrix to be strictly positive definite.

We model the unobservable boundary D as a time invariant random variable which is independent of the driving Brownian motions, and whose distribution *is known to the investor*. We assume that $0 < D < 1$ almost surely. Furthermore, he can observe the initial health $C_0 \geq 1$ and knows that at time $t = 0$ the state of the firm is healthy.

Let F denote the cummulative distribution function of D, i.e.

$$F(x) = \mathbb{P}(D \leq x).$$

For simplicity we also assume that D has a density $f(x)$ which is continuous, strictly positive and bounded on $(0, 1)$.

Formally we model the switch of the firm's state from *healthy* to *distressed* or *default* as the switching of the indicator process

$$N_t \triangleq \mathbb{I}\{t \geq \tau_h\},$$

where

$$\tau_h \triangleq \inf\{t \geq 0 : \ C_t \leq D\}.$$

At any time $t \geq 0$, the investor has full information on the states of the asset prices $S_s^{(1)}, \ldots, S_s^{(n)}$ and the creditworthiness index C_s for $0 \leq s \leq t$. It is natural to assume that the investor also knows whether or not C_t has already hit D, since such an event could be observed in the market. We will therefore assume that at time t the information available to the investor is given by

$$\mathcal{F}_t \triangleq \sigma\left(\left\{B_s^{(1)}, \ \ldots, \ B_s^{(n)} : \ 0 \leq s \leq t\right\} \cup \mathbb{I}\left\{t \geq \tau_h\right\} \cup \mathcal{N}\right).$$

in the default case, and by

$$\mathcal{F}_t \triangleq \sigma\left(\left\{B_s^{(1)}, \ \ldots, \ B_s^{(n)} : \ 0 \leq s \leq t\right\} \cup \mathbb{I}\left\{t \geq \tau_h\right\} \cup \mathbb{I}\left\{t \geq \tau_d\right\} \cup \mathcal{N}\right).$$

in the non-default case. Following Giesecke (2006), the corresponding filtration is called the *investor's filtration*. Note that τ_h obviously is an $\mathcal{F}_t$-stopping time.

We also define the *historic low of* C as

$$m_t \triangleq \min\{C_s : \ 0 \leq s \leq t\}.$$

If at time t the firm is still in a healthy state and if $m_t < 1$, the investor has the additional information $D < m_t$, and hence also gets some information on the distribution on D. We let

$$F_m(x) \triangleq \mathbb{P}(D \leq x \mid D < m)$$

denote the cumulative distribution function of D conditioned on the historic low of C. Obviously,

$$F_m(x) = \begin{cases} \frac{\mathbb{P}(D \leq x)}{\mathbb{P}(D < m)}, & 0 < x < m \\ 1, & x \geq m. \end{cases}$$

For $x < m$, we let

$$f_m(x) \triangleq \frac{dF_m(x)}{dx}$$

be the conditional density of D. Finally, we let

$$f_m(m) \triangleq \lim_{x \to m^-} f_m(x).$$

Given that at time t we have $C_t = m_t = m < 1$, we can consider $f_m(m)$ as the analogue of a default intensity. If in a small time interval $[t, t + \Delta t]$, the historic low decreases from m to $m - \Delta m$, then the probability of the company's health hitting D in $[t, t + \Delta t]$ is approximately $f_m(m)\ \Delta m$.

Since the process N_t is increasing, there exists a $\mathcal{F}_t$-predictable process K_t (the compensator of N_t) such that $N_t - K_t$ is an $\mathcal{F}_t$-martingale (see e.g. Lipster and Shiryaev (2000)). For later we need the following

Lemma 8. *The compensator of N_t is*

$$K_t = \int_0^{\tau_h \wedge t} f_{m_s}(m_s)\ dm_s,$$

i.e.

$$N_t - \int_0^{\tau_h \wedge t} f_{m_s}(m_s)\ dm_s$$

is an $\mathcal{F}_t$-martingale.

For the proof see e.g. Giesecke (2006) and references therein.

As before, we define the value function U at time t as

$$\sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}\left[u(W_T) \mid \mathcal{F}_t\right].$$

This makes it necessary to define the set of admissible trading strategies.

Definition 5. *An admissible trading strategy is an $\mathcal{F}_t$-predictable process $\boldsymbol{\pi}_t = \left(\pi_t^{(1)}, \pi_t^{(2)}\right) \in \mathbb{R}^2$ satisfying the following:*

(i)

$$\pi_t^{(2)} = 0 \quad \textit{for} \quad t > \tau_h,$$

(ii)

$$\int_0^T \boldsymbol{\pi}_t^2 \; dt < \infty \quad \textit{almost surely,}$$

(iii)

$$\mathbb{E} \int_0^T \boldsymbol{\pi}_t^2 \left(e^{-\gamma e^{r(T-t)} W_t^{\boldsymbol{\pi}}} \right)^2 \; dt < \infty.$$

4.2 The Investment Problem

While the optimization problems from the following sections can easily be generalized to the case of n risky assets as described in the previous section, we restrict ourselves to the case of one default-free risky asset I (the stock index) and the defaultable stock S, mainly for the sake of notational simplicity. We first consider the problem in which the investor invests in I, S and the money market only, but not in credit derivatives. As mentioned, we distinguish the non-default and the default case.

4.2.1 The Non-Default Case

For an admissible trading strategy $\boldsymbol{\pi}$, the wealth process has the dynamics

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r \; W_t \right] \; dt + \pi_t^{(1)} \sigma_1 \; dB_t^{(1)} + \pi_t^{(2)} \sigma_2 \; dB_t^{(2)}, & t < \tau_h, \\ \left[(\mu_1 - r) \pi_t^{(1)} + r \; W_t \right] \; dt + \pi_t^{(1)} \sigma_1 dB_t^{(1)}, & t > \tau_h, \end{cases}$$

subject to $W_{\tau_h} = W_{\tau_h^-}$.

Given the above definitions, we define the value function U as

$$U(w, I, S, C, m, t) \triangleq \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E} \left[u(W_T) \mid W_t = w, \; I_t = I, \; S_t = S, \; C_t = C, \; m_t = m, \; \tau_h > t \right].$$

It is easy to see that for any trading strategy, the wealth process is independent of I and S. Hence U is a function of w, C, m, t only and defined on the set

$$\mathcal{D} = \left\{ (w, S, C, t) \in \mathbb{R} \times (0, \infty) \times (0, \infty) \times [0, T] : \; C \geq m \right\}.$$

As long as $m_t > 1$, the observation of the historic low of C does not give the investor any additional information about D. Therefore we intuitively expect the following lemma to hold, showing that it is sufficient to determine U for the case $m \leq 1$.

Lemma 9. *For all $w \in \mathbb{R}$, $C \geq m \geq 1$, $t \in [0, T]$ we have*

$$U(w, C, m, t) = U(w, C, 1, t).$$

Proof. For $m > 1$ we obviously have $f_m(x) = 0$ for $1 < x \leq m$ and $f_m(x) = f_1(x) = f(x)$ for $0 < x \leq 1$. For any admissible trading strategy we have

$$\begin{aligned}
&\mathbb{E}_t\left[u(W_t) \mid W_t = w,\ C_t = C,\ m_t = m,\ t < \tau_h\ \right] \\
&\qquad = \int_0^m \mathbb{E}_t\left[u(W_t) \mid W_t = w,\ C_t = C,\ m_t = m,\ D = \xi\ \right]\ f_m(\xi)\ d\xi \\
&\qquad = \int_0^1 \mathbb{E}_t\left[u(W_t) \mid W_t = w,\ C_t = C,\ m_t = m,\ D = \xi\ \right]\ f_1(\xi)\ d\xi \\
&\qquad\qquad = \mathbb{E}_t\left[u(W_t) \mid W_t = w,\ C_t = C,\ m_t = 1,\ t < \tau_h\ \right]. \qquad (4.1)
\end{aligned}$$

The lemma follows from taking the supremum over all admissible trading strategies. □

In the following we determine the HJB equation for U when $0 < m \leq 1$. In principle it would be sufficient to refer to the result in the verification theorem in section 4.A. However, since the resulting HJB equation is somewhat non-standard, we derive it heuristically. Suppose that $W_{t+\Delta t} = W + \Delta W$, $C_{t+\Delta t} = C + \Delta C$, $m_{t+\Delta t} = m + \Delta m$ for some small time interval $[t, t + \Delta t]$. Note that with this notation, obviously $\Delta m < 0$. Furthermore, let $\tau \triangleq \tau_h \wedge (t + \Delta t)$.

We start with the case $C > m$ and define the stopping time $\tau \triangleq (t + \Delta t) \cap \inf\{\bar{t} \geq t :\ C_{\bar{t}} = m\}$. Then there is no default in the time interval $[t, \tau]$, and following Bellman's principle of optimality, we expect the following to hold for any trading strategy $\boldsymbol{\pi}$:

$$U(w, C, m, t) \geq \mathbb{E}_t[U(W_\tau, C_\tau, m_\tau, \tau)].$$

Furthermore, we expect to have equality for the optimal strategy $\boldsymbol{\pi}^*$.

The evolution of U is given by

$$dU = (\partial_t U + \mathcal{L}^{\boldsymbol{\pi}} U)\ dt + \pi^{(1)}\sigma_1\ \partial_w U\ dB_t^{(1)} + \pi^{(2)}\sigma_2\ \partial_w U\ dB_t^{(2)} + \\ + \eta C\ \partial_C U\ dB_t^{(3)} + \partial_m U\ dm_t$$

with

$$\mathcal{L}^{\boldsymbol{\pi}} U = (rw + \boldsymbol{\pi}^T(\boldsymbol{\mu} - \boldsymbol{r}))\partial_w U + \nu C \partial_C U + \frac{1}{2}\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi}\ \partial_{ww} U + \frac{1}{2}\eta^2 C^2\ \partial_{CC} U + \boldsymbol{\pi}^T \boldsymbol{\omega} C\ \partial_{wC} U.$$

Note that since m_t is a decreasing process and hence of bounded variation, there are no second order derivatives containing m. From Ito's lemma we therefore get

$$\mathbb{E}_t[U(W_\tau, C_\tau, m_\tau, \tau)] = U(w, C, m, t) + \mathbb{E}_t\left[\int_t^\tau (\partial_t U + \mathcal{L}^{\boldsymbol{\pi}} U)\ ds\right] + \mathbb{E}_t\left[\int_t^\tau \partial_m U\ dm_s\right].$$

For $s \in [t, \tau]$, we have $dm_s = 0$, because the historic low of C does not decrease in this interval. We therefore get the inequality

$$0 \geq \mathbb{E}_t\left[\int_t^\tau (\partial_t U + \mathcal{L}^{\boldsymbol{\pi}} U)\ ds\right] \\ = \mathbb{E}_t\left[\int_t^{t+\Delta t} (\partial_t U + \mathcal{L}^{\boldsymbol{\pi}} U)\ ds \cdot \mathbb{I}\{\tau = t + \Delta t\} + \int_t^\tau (\partial_t U + \mathcal{L}^{\boldsymbol{\pi}} U)\ ds \cdot \mathbb{I}\{\tau < t + \Delta t\}\right].$$

We divide this inequality by Δt. Since it can easily be seen that

$$\frac{1}{\Delta t}\mathbb{P}(\tau \leq t + \Delta t) \to 0 \quad (\Delta t \to 0),$$

the second term on the right hand side (after division by Δt) approaches 0. Therefore we expect

$$\partial_t U + \mathcal{L}^{\boldsymbol{\pi}} U \leq 0$$

to hold for any admissible trading strategy $\boldsymbol{\pi}$. For the optimal strategy $\boldsymbol{\pi}^*$ we expect to get equality, so the HJB equation becomes

$$\partial_t U + \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} U = 0.$$

Now we consider the case $C = m$ and define the sequence of stopping times

$$\tau_k = \inf\{t :\ m_t = m - \frac{1}{k}\} \wedge T, \quad k = 1,\ 2,\ 3,\ \ldots.$$

By a generalization of Bellman's principle, for any admissible trading strategy we expect to have the inequality

$$U(w,C,m,t) \geq \mathbb{E}_t \left[U(W+\Delta W,\ C+\Delta C,\ m+\Delta m,\ t+\Delta t) \cdot \mathbb{I}\left\{\tau_h > t+\Delta t\right\} + \right.$$
$$\left. + V(w_{\tau_h}, \tau_h) \cdot \mathbb{I}\{t < \tau_h \leq t+\Delta t\} \right].$$

with equality for the optimal strategy $\boldsymbol{\pi}^*$.

Working under $\boldsymbol{\pi}^*$ and applying Ito's lemma to U as before therefore yields the equation

$$U(w,C,m,t) = \mathbb{E}_t \left[\left\{ U(w,C,m,t) + \int_t^{\tau_k} \left(\partial_t U + \mathcal{L}^{\boldsymbol{\pi}^*} U\right)\ ds + \int_t^{\tau_k} \partial_m U\ dm_s \right\} \cdot \mathbb{I}\{\tau_h > \tau_k\} \right] +$$
$$+ \mathbb{E}_t \left[V(W_{\tau_h}, \tau_h) \cdot \mathbb{I}\{\tau_k < \tau_h\} \right].$$

We already know that if $C_s > m_s$, we can expect that $\partial_t U + \mathcal{L}^{\boldsymbol{\pi}^*} U = 0$, while the set $\{s:\ C_s = m_s\}$ is of measure 0, and therefore the Riemann integral over this set does not contribute anything. . Therefore the equation above yields

$$\mathbb{E}_t \left[\int_t^{\tau_k} \partial_m U\ dm_s \cdot \mathbb{I}\{\tau_h > \tau_k\} \right] = \mathbb{E}_t \left[(U(w,C,m,t) - V(W_{\tau_h}, \tau_h)) \cdot \mathbb{I}\{\tau_h \leq \tau_k\} \right].$$

Now we multiply this equation by k (equivalently, we divide by $\frac{1}{k}$). By the definition of τ_k and $f_m(m)$, we get

$$k\ \mathbb{P}(\tau_h \leq \tau_k \mid m_t = m) = k\ \mathbb{P}\left(m - \frac{1}{k} \leq D < m \right) \to f_m(m) \quad (k \to \infty).$$

Since $\mathbb{P}(\tau_k > \tau_h) \to 0$ $(k \to 0)$, and since $k = \frac{1}{m - m_{\tau_k}}$ if $\tau_k < \tau_h$, we expect the left hand side of (4.2.1) to converge to $\partial_m U$ (after division by $\frac{1}{k}$), while the right hand side is expected to approach $f_m(m)\ \ [U(w,C,m,t) - V(w,t)]$. Consequently, the full HJB equation for U reads

$$\begin{cases} \partial_t U + \sup\limits_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} U = 0 \\ \quad U(w,C,m,T) = u(w), \quad w \in \mathbb{R},\ C \geq m, \\ \qquad\qquad \partial_m U = f_m(m)\ \left[U(w,C,m,t) - V(w,t)\right], \quad C = m. \end{cases} \tag{4.2}$$

As in the case with visible barrier, the optimal strategy is given by

$$\boldsymbol{\pi}^* = -\frac{1}{\partial_{ww}U}\ \boldsymbol{\Omega}^{-1}\left[(\boldsymbol{\mu}-\boldsymbol{r})\ \partial_w U + \boldsymbol{\omega} C\ \partial_{wC}U\right]. \tag{4.3}$$

It should be pointed out that this equation can also be obtained through a martingale argument by using lemma 8. Under the optimal strategy we expect the process

$$M_t^{(1)} = \begin{cases} U(W_t,\ C_t,\ m_t,\ t), & t < \tau_h, \\ V(W_t,\ t), & t \geq \tau_h \end{cases}$$

to be a martingale. If we let

$$\Delta_t = V(W_t, t) - U(W_t,\ C_t,\ m_t,\ t),$$

then by lemma 8, the process

$$M_t^{(2)} = \Delta_{\tau_h} \cdot \mathbb{I}\{t < \tau_h\} - \int_0^{t\wedge\tau_h} \Delta_s\ dK_s$$

is also a martingale, so that $M_t^{(1)} - M_t^{(2)}$ is a predictable martingale. Applying Ito's lemma then yields the PDE and the boundary condition.

To simplify equation (4.2) we proceed as in the case for the visible barrier. We make the ansatz $U(w, C, m, t) = u(we^{r(T-t)})\ e^{-\frac{1}{2}\Lambda(T-t)}\ G^{\beta}\left(\ln C,\ \ln m,\ T-t\right)$ with

$$\beta = \frac{1}{1 - \frac{1}{\eta^2}\boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega}}, \quad \tilde{\nu} = \nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu}-\boldsymbol{r})^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega}, \tag{4.4}$$

which yield the following heat-like equation for $G(x, y, \tau)$:

$$\begin{cases} -\partial_\tau G + \tilde{\nu}\ \partial_x G + \dfrac{1}{2}\eta^2\ \partial_{xx}G = 0, \\ \qquad\qquad G(x, y, 0) = 1, \quad x \geq y, \\ \qquad\qquad\qquad \partial_y G = f_m(m)\ m\ \dfrac{1}{\beta}\ \left[G - e^{\frac{1}{2}(\Lambda^2-\lambda^2)(T-t)}\ G^{1-\beta}\right], \quad x = y. \end{cases} \tag{4.5}$$

In the last line above we could of course replace m by e^y. Note that (4.5) is not a two-dimensional heat equation, because the derivatives with respect to y are missing. It is

however a family of one-dimensional heat equations in x and τ which are coupled by the second boundary condition.

In terms of G, the optimal trading strategy is given by

$$\boldsymbol{\pi}^* = \frac{1}{\gamma e^{r(T-t)}}\ \boldsymbol{\Omega}^{-1}\ \left[(\boldsymbol{\mu}-\boldsymbol{r}) + \beta\boldsymbol{\omega}\ \frac{\partial_x G}{G}\right].$$

To conclude this section, we prove the following lemma, since its result is crucial for future sections:

Lemma 10. *Under the assumptions of this section and β defined as in equation (4.4), we have*

$$\beta > 1.$$

Proof. It suffices to show that $0 < \boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega} < \eta^2$. By the result in appendix 1.A,

$$\left(\begin{array}{c|c} \boldsymbol{\Omega} & \boldsymbol{\omega} \\ \hline \boldsymbol{\omega}^T & \eta^2 \end{array}\right)^{-1} = \left(\begin{array}{c|c} (\boldsymbol{\Omega}-\boldsymbol{\omega}\eta^{-2}\boldsymbol{\omega}^T)^{-1} & -\boldsymbol{\Omega}^{-1}\boldsymbol{\omega}\left(-\boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega}+\eta^2\right)^{-1} \\ \hline -\left(-\boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega}+\eta^2\right)^{-1}\boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1} & \left(\eta^2-\boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega}\right)^{-1} \end{array}\right).$$

Since the matrix on the left hand side is positive definite, the same applies for the matrix on the right hand side. Consequently, the bottom right entry is positive, i.e. $\eta^2 - \boldsymbol{\omega}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\omega} > 0$. Since both terms in this difference are positive, this proves the lemma. □

4.2.2 The Default Case

The heuristic derivation of the HJB equation is very similar to the one in section 4.2.1. The only difference is that at τ_h, the investors wealth changes from $W_{\tau_h^-}$ to $W_{\tau_h^-} - \pi^{(2),*}_{\tau_h}$. Here $\pi^{(2),*}_t$ is the second component of the optimal trading strategy $\boldsymbol{\pi}^*$, i.e. the dollar amount invested in S at time t.

Under this model, the wealth process satisfies

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu}-\boldsymbol{r})^T\boldsymbol{\pi}_t + r\ W_t\right]\ dt + \pi^{(1)}_t\sigma_1\ dB^{(1)}_t + \pi^{(2)}_t\sigma_2\ dB^{(2)}_t, & t < \tau_h, \\ \left[(\mu_1 - r)\pi^{(1)}_t + r\ W_t\right]\ dt + \pi^{(1)}_t\sigma_1 dB^{(1)}_t, & t > \tau_h, \end{cases}$$

subject to $W_\tau = W_{\tau^-} - \pi_\tau^{(2),*} \cdot \mathbb{I}\{\tau_h < T\}$.

Consequently, U also satisfies the HJB equation (4.2), except that the second boundary condition has to be modified.

The main result of this section and the accompanying verification theorem in section 4.A is the fact that on the boundary $C = m$, it is optimal for the investor to be as short as possible in S. This may seem somewhat counterintuitive, since a non-default of S may be a disadvantage for the investor. However, the expected change in wealth caused by a change of the driving Brownian motions during a short time interval of length dt is proportional to dt, and for every path the set $\{t : \ C_t = m_t\}$ is of measure 0. In contrast, if $C_t = m_t$ the expected change in wealth caused by the default of S in an interval of length dt is proportional to $\pi^{(2),*} dm_t$.

For this reason we have to make the following restrictions on $\mathcal{A}$:

Assumption 11. *There exists a constant c such that every admissible trading strategy $\boldsymbol{\pi} \in \mathcal{A}$ has the property*

$$\pi_t^{(2)} \geq -c \quad \textit{for all } t \in [0, T] \textit{ almost surely.}$$

Then the full HJB equation for U reads

$$\begin{cases} \partial_t U + \sup\limits_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} U = 0 \\ \quad U(w, C, m, T) = u(w) \\ \qquad\qquad \partial_m U = f_m(m) \ \left[U(w, C, m, t) - V(w + c, t)\right]. \end{cases} \tag{4.6}$$

This equation is the same as for the non-default case, except for the term $V(w + c, t)$ in the boundary condition at $C = m$.

As in the non-default case we can simplify this equation via the substitution $U(w, C, m, t) = u(we^{r(T-t)})\ e^{-\frac{1}{2}\Lambda(T-t)}\ G^\beta\left(\ln C,\ \ln m,\ T - t\right)$ with β and Λ as before to get the following

equation for $G(x, y, \tau)$:

$$\begin{cases} -\partial_\tau G + \tilde{\nu}\ \partial_x G + \dfrac{1}{2}\eta^2\ \partial_{xx} G = 0, \\ \qquad\qquad G(x, y, 0) = 1, \\ \qquad\qquad\qquad \partial_y G = f_m(m)\ m\ \dfrac{1}{\beta}\ \left[G - e^{\frac{1}{2}(\Lambda^2-\lambda^2)(T-t)}\ G^{1-\beta}\right], \quad x = y. \end{cases} \tag{4.7}$$

4.3 Reduction of Dimension

We now consider the case when D has the initial distribution

$$F(x) = \begin{cases} x^k, & 0 \le x < 1, \\ 1, & x \ge 1 \end{cases}$$

for some constant $k \ge 0$. For $k = 0$ this includes the case that D is uniformly distributed in $(0, 1)$. One easily finds that for $0 < m < 1$,

$$\mathbb{P}(D \le x \mid D < m) = \begin{cases} \frac{\mathbb{P}(D \le x)}{\mathbb{P}(D < m)} = \frac{x^k}{m^k}, & 0 \le x < m, \\ 1, & x \ge m, \end{cases}$$

and hence one easily sees that

$$f_m(x) = k \cdot \frac{x^{k-1}}{m^k} \cdot \mathbb{I}\{0 < x < m\}$$

and $f_m(m) = \frac{k}{m}$.

We show the following

Lemma 12. *For $0 < m < 1$, U depends on w, $\frac{C}{m}$, t only, i.e. $U(w, C, m, t) = U_0(w, \frac{C}{m}, t)$ for some function U_0.*

Proof. Fix $t \in [0, T]$ and let $w \in \mathbb{R}$, $0 < m^{(1)}, m^{(2)} < 1$, $C^{(1)}$, $C^{(2)}$ be given such that $\frac{C^{(1)}}{m^{(1)}} = \frac{C^{(2)}}{m^{(2)}}$. We have to show that

$$U(w, C^{(1)}, m^{(1)}, t) = U(w, C^{(2)}, m^{(2)}, t).$$

Let $\alpha \triangleq \frac{m^{(2)}}{m^{(1)}} = \frac{C^{(2)}}{C^{(1)}}$. For any admissible strategy $\boldsymbol{\pi}$ we have

$$\mathbb{E}_t\left[u(W_T) \;\middle|\; W_t = w,\ C_t = C^{(1)},\ m_t = m^{(1)}, t < \tau_h\right] = \int_0^{m^{(1)}} \mathbb{E}_t\left[u(W_T) \;\middle|\; W_t = w,\ C_t = C^{(1)},\ m_t = m^{(1)}, D = \xi\right]\ f_{m^{(1)}}(\xi)\ d\xi, \quad (4.8)$$

and the analogous formula holds for the superscript (1) replaced by (2). Given the dynamics of C_s, we can rewrite the right hand side of (4.8) as

$$\int_0^{m^{(1)}} \mathbb{E}_t\left[u(W_T) \;\middle|\; W_t = w,\ C_t = C^{(2)},\ m_t = m^{(2)}, D = \alpha\xi\right]\ f_{m^{(1)}}(\xi)\ d\xi$$

To manipulate this term, we make a change of variables $\zeta = \alpha\xi$. We get

$$\begin{aligned} f_{m^{(1)}}(\xi)\ d\xi &= k \cdot \frac{\xi^{k-1}}{\left(m^{(1)}\right)^k}\ d\xi \\ &= k \cdot \frac{\alpha\zeta^{k-1}}{\left(m^{(2)}\right)^k} \cdot \frac{1}{\alpha}\ d\zeta \\ &= k \cdot \frac{\zeta^{k-1}}{\left(m^{(2)}\right)^k}\ d\zeta = f_{m^{(2)}}(\zeta)\ d\zeta. \end{aligned}$$

Therefore the right hand side of (4.8) becomes

$$\int_0^{m^{(2)}} \mathbb{E}_t\left[u(W_T) \;\middle|\; W_t = w,\ C_t = C^{(2)},\ m_t = m^{(2)}, D = \zeta\right]\ f_{m^{(2)}}(\zeta)\ d\zeta$$

which equals

$$\mathbb{E}_t\left[u(W_T) \;\middle|\; W_t = w,\ C_t = C^{(2)},\ m_t = m^{(2)}, t < \tau_h\right].$$

Taking the supremum over all trading strategies proves the lemma. □

If the assumption on the distribution of D from this section is satisfied, then the function G from the substitution $U(w, C, m, t) = u(we^{r(T-t)})\, e^{-\frac{1}{2}\Lambda^2(T-t)} G^{\beta}(\ln C,\ \ln m, T-t)$ from the previous section depends on $x - y$ and τ only, i.e. $G(x, y, \tau) = G_0(x - y, \tau)$, where $G(z, \tau)$ is a function defined on $[0, \infty) \times [0, T]$.

Using the equations for G from the previous section, it follows that the function $G_0(z, \tau)$ satisfies a one-dimensional heat equation. For the *non-default* case, this equation

reads

$$\begin{cases} -\partial_\tau G_0 + \tilde{\nu}\ \partial_z G_0 + \dfrac{1}{2}\eta^2\ \partial_{zz} G_0 = 0 \\ \qquad\qquad G_0(z,0) = 1, \quad z > 0, \\ \qquad\qquad \partial_z G_0 = \dfrac{1}{\beta}\ \left[e^{\frac{1}{2}(\Lambda^2-\lambda^2)\tau}\ G_0^{1-\beta} - G_0\right], \quad z = 0, \end{cases} \tag{4.9}$$

whereas for the *default case* we get

$$\begin{cases} -\partial_\tau G_0 + \tilde{\nu}\ \partial_z G_0 + \dfrac{1}{2}\eta^2\ \partial_{zz} G_0 = 0 \\ \qquad\qquad G_0(z,0) = 1, \quad z > 0, \\ \qquad\qquad \partial_z G_0 = \dfrac{1}{\beta}\ \left[e^{\frac{1}{2}(\Lambda^2-\lambda^2)\tau}\ G_0^{1-\beta}\ e^{a_t c} - G_0\right], \quad z = 0. \end{cases} \tag{4.10}$$

4.4 Pricing of Credit Derivatives

As in previous sections we define the value functions corresponding to investments in defaultable bonds and credit default swaps and derive the corresponding HJB-equations. Because of their similarity to the ones in sections 4.2.1 and 4.2.2 we treat the bond and the CDS simultaneously.

4.4.1 The Non-Default Case

Let $\tau_1 = \tau_h \wedge T$, $\tau_2 = \tau_d \wedge T$. We assume that the wealth process has the dynamics

$$d\overline{W}_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ \overline{W}_t + \epsilon A F\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t < \tau_1, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \overline{W}_t + \epsilon A F\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_1 < t < \tau_2, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_2 < t < T \end{cases}$$

subject to $\overline{W}_{\tau_1} = \overline{W}_{\tau_1^-} + R_2 \cdot \mathbb{I}\{\tau_1 = T\}$, $\overline{W}_{\tau_2} = \overline{W}_{\tau_2^-} + R_1 \cdot \mathbb{I}\{\tau_2 < T\} + R_2 \cdot \{\tau_2 = T\}$. The choice $\epsilon = 0$, $R_1 = RF$, $R_2 = F$ corresponds to the defaultable bond, whereas the values $\epsilon = \pm 1$, $R_1 = \epsilon(1-R)F$, $R_2 = 0$ correspond to the credit default swap.

We define

$$\overline{U}(w, C, m, t) \triangleq \sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}\left[u(\overline{W}_T) \mid \overline{W}_t = w,\ C_t = C,\ m_t = m,\ t < \tau_h\right].$$

Then the corresponding HJB equation is

$$\begin{cases} \partial_t \overline{U} + \sup_{\boldsymbol{\pi}\in\mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} \overline{U} = 0 \\ \overline{U}(w, C, m, T) = u(w + R_2) \\ \partial_m \overline{U} = f_m(m) \ \left[\overline{U}(w, C, m, t) - \overline{V}(w, t)\right], \end{cases} \tag{4.11}$$

where $\overline{V}$ is the value function

$$\overline{V}(w, t) = \sup_{\pi\in\mathbb{R}} \mathbb{E}[u(W_T) \mid W_t = w, \ t \geq \tau_h].$$

For the defaultable bond, $\overline{V}$ is the same function as in section 2.1.1, while for the credit default swap, $\overline{V}$ from the equation above is the same as $\tilde{V}$ from from section 2.1.2.

The substitution $U(w, C, m, t) = u(we^{r(T-t)}) \ e^{-\frac{1}{2}\Lambda(T-t)} \ e^{\Psi(T-t)} \ \overline{G}^{\beta} \ (\ln C, \ \ln m, \ T - t)$ with

$$\Psi(\tau) = -\epsilon\gamma \frac{A}{r} e^{r\tau}$$

then yields the following equation for $\overline{G}(x, y, \tau)$:

$$\begin{cases} -\partial_\tau \overline{G} + \tilde{\nu} \ \partial_x \overline{G} + \frac{1}{2}\eta^2 \ \partial_{xx} \overline{G} = 0, \\ \overline{G}(x, y, 0) = e^{-\frac{\gamma R_2}{\beta}}, \\ \partial_y \overline{G} = f_m(m) \ m \ \frac{1}{\beta} \ \left[\overline{G} - e^{\frac{1}{2}\Lambda^2\tau} \ \overline{G}^{1-\beta} \ \overline{g}(T - \tau)\right], \quad x = y. \end{cases} \tag{4.12}$$

Again, $\overline{g}$ is the same as in section 2.1.1 for the bond, and $\tilde{g}$ from section 2.1.2 for the CDS.

For the short-term credit spreads we expect the following qualitative behaviour:

- If $C_t > m_t = m$ close to maturity, then survival of the firm can be anticipated, and hence short-term spreads should tend to 0. We expect them to go to 0 even faster as if m were a visible critical level (i.e. $D = m$ in chapter 3), because here upon C_t hitting m, the reference entity may switch to the distressed regime soon after, but does not have to.

- If $C_t = m_t$, switching may occur in the next instant, and hence we expect to get a positive credit spread, even for short maturities.

In fact, in the risk-neutral analogue, Giesecke (2006) finds that credit spreads for $C_t = m_t$ approach ∞ as the time to maturity approaches 0. It would be interesting to see whether this is the case here as well.

4.4.2 The Default Case

In this case we let $\tau = \tau_h \wedge T$ and assume that the wealth process has the dynamics

$$d\overline{W}_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ \overline{W}_t + \epsilon AF\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t < \tau, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau < t < T \end{cases}$$

subject to $\overline{W}_\tau = \overline{W}_{\tau^-} + \left(-\pi_\tau^{(2)} + R_1\right) \cdot \mathbb{I}\{\tau < T\} + R_2 \cdot \mathbb{I}\{\tau = T\}$ where we get the cases of the defaultable bond and the CDS with the same choices for ϵ, R_1, R_2 as in the non-default case. The HJB equation for U is very similar to the one in the non-default case. However there is a difference due to the fact that the investor makes/receives a random recovery payment at time τ_h, which is different compared to the non-default case.

We let

$$\tilde{R}_t = -\frac{1}{\gamma e^{r(T-t)}} \log \mathbb{E}\ e^{-\gamma R_1 e^{r(T-t)}}, \quad \text{i.e.} \quad e^{-\gamma \tilde{R}_t e^{r(T-t)}} = \mathbb{E}\ e^{-\gamma R_1 e^{r(T-t)}}.$$

The the HJB equation for $\tilde{U}$ is

$$\begin{cases} \partial_t \overline{U} + \epsilon AF\ \partial_w \overline{U} + \sup\limits_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} \overline{U} = 0 \\ \qquad U(w, C, m, T) = u(w), \quad w \in \mathbb{R}, \\ \qquad \partial_m \overline{U} = f_m(m)\ \Big[\overline{U}(w, C, m, t) - \\ \qquad\qquad -V(w + c, t) \cdot e^{-\gamma \tilde{R}_t e^{r(T-t)}}\Big], \quad C = m. \end{cases} \tag{4.13}$$

Here V is the value function for the standard Merton problem with investment in I and the money market account only.

The same substitution then again yields a heat-like equation for $\overline{G}(x, y, \tau)$:

$$\begin{cases} -\partial_\tau \overline{G} + \tilde{\nu}\ \partial_x \overline{G} + \dfrac{1}{2}\eta^2\ \partial_{xx}\overline{G} = 0, \\ \overline{G}(x, y, 0) = e^{-\frac{\gamma R_2}{\beta}}, \\ \partial_y \overline{G} = f_m(m)\ m\ \dfrac{1}{\beta}\ \left[\overline{G} - e^{\frac{1}{2}(\Lambda^2-\lambda^2)\tau}\ e^{a_t c}\ \overline{G}^{1-\beta}\right], \quad x = y. \end{cases} \tag{4.14}$$

4.A Appendix: Verification Theorems

In this appendix we show that the solutions of the HJB equations from this chapter are indeed the value functions of the corresponding optimization problems. We treat the investment problem, the bond and the credit default swap simultaneously, but we separate the non-default and the default case.

4.A.1 The Non-Default Case

As in section 4.4.1, let $\tau_1 = \tau_h \wedge T$, $\tau_2 = \tau_d \wedge T$ and assume that the wealth process has the dynamics

$$dW_t = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ W_t + \epsilon A F\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t < \tau_1, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ W_t + \epsilon A F\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_1 < t < \tau_2, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ W_t\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_2 < t < T \end{cases}$$

subject to $W_{\tau_1} = W_{\tau_1^-} + R_2 \cdot \mathbb{I}\{\tau_1 = T\}$, $W_{\tau_2} = W_{\tau_2^-} + R_1 \cdot \mathbb{I}\{\tau_2 < T\} + R_2 \cdot \{\tau_2 = T\}$. Consider the HJB equation,

$$\begin{cases} \partial_t \overline{U} + \epsilon A \partial_x \overline{U} + \sup\limits_{\boldsymbol{\pi}\in\mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} \overline{U} = 0 \\ \overline{U}(w, C, m, T) = u(w) \\ \partial_m \overline{U} = f_m(m)\ \left[\overline{U}(w, C, m, t) - \overline{V}(w, t)\right], \quad C = m. \end{cases}$$

Let $\overline{\pi}$ be the optimal strategy corresponding to $\overline{V}$. Then the following verification theorem holds:

Theorem 13. *Suppose that there is a function $H(w, C, m, t)$ which is a solution of equation (4.2) for $w \in \mathbb{R}$, $t \in [0, T]$, $C \geq m$, $0 < m \leq 1$and which is sufficiently integrable. Furthermore, suppose that for each (w, C, m, t) satisfying $w \in \mathbb{R}$, $t \in [0, T]$, $C \geq m$, $0 < m \leq 1$ there exists $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(w, C, m, t) \in \mathbb{R}^n$ such that*

$$\mathcal{L}^{\boldsymbol{\pi}^*} H = \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} H \tag{4.15}$$

and such that the trading strategy

$$\overline{\boldsymbol{\pi}}_t = \begin{cases} \boldsymbol{\pi}^*(W_t, C_t, m_t, t), & t < \tau_h,\ C_t > m_t, \\ \textit{arbitrary}, & t < \tau_h,\ C_t = m_t, \\ \overline{\pi}_t, & t \geq \tau_h \end{cases}$$

is admissible. Then $U = H$ and $\overline{\boldsymbol{\pi}}$ is an optimal strategy.

Proof. We define the stopping time $\tau = \tau_h \wedge T$. Since $\overline{W}_t$ is an Ito diffusion in $[0, \tau]$ for every admissible trading strategy $\boldsymbol{\pi}$, we have

$$\mathbb{E}_t \left[H(W_\tau, C_\tau, m_\tau, \tau) \right] = U(w, C, m, t) + \mathbb{E}_t \left[\int_t^\tau \left(\partial_t H + \mathcal{L}^{\boldsymbol{\pi}} H \right) \, ds \right] + \mathbb{E}_t \left[\int_t^\tau \partial_m H \, dm_s \right].$$

Since H satisfies the HJB equation, we get the inequality

$$\begin{aligned} \mathbb{E}_t \left[H(W_\tau, C_\tau, m_\tau, \tau) \right] \geq & U(w, C, m, t) + \\ & + \mathbb{E}_t \left[\int_t^\tau f_{m_s}(m_s) \, \left[H(W_s^{(c)}, C_s, m_s, s) - V(W_s^{(c)}, s) \right] \, dm_s \right]. \end{aligned}$$

Note that above we could replace $\partial_m H$ by the last term on the right hand side, because if $C_s > m_s$, then $dm_s = 0$, while H satisfies the corresponding boundary condition, if $C_s = m_s$. Hence we get

$$\begin{aligned} H(w, C, m, t) \geq & \mathbb{E}_t \left[H(W_\tau, C_\tau, m_\tau, \tau) \cdot \mathbb{I}\{\tau = T\} \right] + \mathbb{E}_t \left[H(W_\tau, C_\tau, m_\tau, \tau) \cdot \mathbb{I}\{\tau < T\} \right] - \\ & - \mathbb{E}_t \left[\int_t^\tau f_{m_s}(m_s) \, \left[H(W_s, C_s, m_s, s) - \overline{V}(W_s, s) \right] \, dm_s \right]. \end{aligned}$$

Since

$$N_t - \int_0^{\tau_h \wedge t} f_{m_s}(m_s) \, dm_s$$

is an $\mathcal{F}_t$-martingale, we get for the last term on the right hand side

$$\begin{aligned}
&\mathbb{E}_t\left[\int_t^\tau f_{m_s}(m_s)\ \left[H(W_s,C_s,m_s,s)-\overline{V}(W_s,s)\right]\ dm_s\right]\\
&=\mathbb{E}_t\left[\int_t^\tau \left[H(W_s,C_s,m_s,s)-\overline{V}(W_s,s)\right]\ dN_s\right]\\
&=\mathbb{E}_t\left[\left[H(W_\tau,C_\tau,m_\tau,\tau)-\overline{V}(W_\tau,\tau)\right]\cdot\mathbb{I}\left\{\tau<T\right\}\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
H(w,C,m,t)&\geq \mathbb{E}_t\left[H(W_T,\ C_T,\ m_T,\ T)\cdot\mathbb{I}\left\{\tau=T\right\}\right]+\mathbb{E}_t\left[\overline{V}(W_\tau,\tau)\cdot\mathbb{I}\{\tau<T\}\right]\\
&=\mathbb{E}_t\left[H(W_T-R_2,\ C_T,\ m_T,\ T)\cdot\mathbb{I}\{\tau=T\}\right]+\mathbb{E}_t\left[\overline{V}(W_\tau,\tau)\cdot\mathbb{I}\{\tau<T\}\right]\\
&\geq \mathbb{E}_t\left[u(W_T)\cdot\mathbb{I}\{\tau=T\}\right]+\mathbb{E}_t\left[u(W_T)\cdot\mathbb{I}\{\tau<T\}\right]\\
&=\mathbb{E}_t\left[u(W_T)\right].
\end{aligned}$$

Since this inequality holds for any admissible strategy, we get $H\geq U$. On the other hand we have equality for $\boldsymbol{\pi}=\overline{\boldsymbol{\pi}}$, so $H=U$.

□

4.A.2 The Default Case

As in this section 4.4.2, we let $\tau=\tau_h\wedge T$ and assume that the wealth process has the dynamics

$$dW_t=\begin{cases}\left[(\boldsymbol{\mu}-\boldsymbol{r})^T\boldsymbol{\pi}_t+r\ W_t+\epsilon AF\right]\ dt+\pi_t^{(1)}\sigma_1\ dB_t^{(1)}+\pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t<\tau,\\ \left[(\mu_1-r)\pi_t^{(1)}+r\ W_t\right]\ dt+\pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau<t<T\end{cases}$$

subject to $W_\tau=W_{\tau^-}+\left(-\pi_\tau^{(2)}+R_1\right)\cdot\mathbb{I}\{\tau<T\}+R_2\cdot\mathbb{I}\{\tau=T\}$.

We consider the corresponding HJB equation

$$\begin{aligned}
&\partial_t U+\sup_{\boldsymbol{\pi}\in\mathbb{R}^2}\mathcal{L}^{\boldsymbol{\pi}}U=0\\
&U(w,C,m,t)=u(w+R_2),\quad w\in\mathbb{R},\ C\geq m,\ 0<m\leq 1,\\
&\partial_m U=f_m(m)\ [U-V(w+c,t)\cdot e^{-\gamma\tilde{R}_t e^{r(T-t)}}],\quad C=m.
\end{aligned}$$

Then the following verification theorem holds:

Theorem 14. *Suppose that there is a function $H(w, C, m, t)$ which is a solution of equation (4.A.2) for $w \in \mathbb{R}$, $t \in [0, T]$, $C \geq m$, $0 < m \leq 1$and which is sufficiently integrable. Furthermore, suppose that for each (w, C, m, t) satisfying $w \in \mathbb{R}$, $t \in [0, T]$, $C \geq m$, $0 < m \leq 1$ there exists $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(w, C, m, t) \in \mathbb{R}^2$ such that*

$$\mathcal{L}^{\boldsymbol{\pi}^*} H = \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \mathcal{L}^{\boldsymbol{\pi}} H \tag{4.16}$$

and such that the trading strategy

$$\overline{\boldsymbol{\pi}}_t = \begin{cases} \boldsymbol{\pi}^*(W_t, C_t, m_t, t), & t < \tau_h,\ C_t > m_t, \\ -c, & t \leq \tau_h,\ C_t = m_t, \\ \pi_t^M, & t > \tau_h \end{cases}$$

is admissible. Then $U = H$ and $\overline{\boldsymbol{\pi}}$ is an optimal strategy.

Proof. We fix t, and as in the verification theorem for the non-default case we consider the continuous part $W_t^{(c)}$ of the wealth process, which follows the dynamics

$$d\overline{W}_s = \begin{cases} \left[(\boldsymbol{\mu} - \boldsymbol{r})^T \boldsymbol{\pi}_t + r\ \overline{W}_s + \epsilon AF\right]\ ds + \pi_t^{(1)} \sigma_1\ dB_s^{(1)} + \pi_t^{(2)} \sigma_2\ dB_s^{(2)}, & s < \tau, \\ \left[(\mu_1 - r)\pi_t^{(1)} + r\ \overline{W}_t\right]\ ds + \pi_t^{(1)} \sigma_1 dB_s^{(1)}, & \tau < s < T \end{cases}$$

subject to $W_\tau^{(c)} = W_{\tau^-}^{(c)}$.

Following the steps from the non-default case, we immediately see that

$$\begin{aligned} H(w, C, m, t) \geq & \mathbb{E}_t \left[H(W_T^{(c)},\ C_T,\ m_T,\ T) \cdot \mathbb{I}\left\{\tau = T\right\}\right] + \\ & + \mathbb{E}_t \left[V(W_\tau^{(c)} + c,\ \tau) \cdot e^{-\gamma \tilde{R}_\tau e^{r(T-\tau)}} \cdot \mathbb{I}\{\tau < T\}\right]. \end{aligned}$$

As before, we have $W_T^{(c)} \cdot \mathbb{I}\{\tau = T\} = (W_T - R_2) \cdot \mathbb{I}\{\tau = T\}$ and hence

$$\mathbb{E}_t \left[H(W_T^{(c)},\ C_T,\ m_T,\ T) \cdot \mathbb{I}\left\{\tau = T\right\}\right] = \mathbb{E}_t \left[u(W_T) \cdot \mathbb{I}\left\{\tau = T\right\}\right].$$

For the second term on the right hand side, we get

$$
\begin{aligned}
&\mathbb{E}_t\left[V(W_\tau^{(c)}+c,\ \tau)\cdot e^{-\gamma\tilde{R}_\tau e^{r(T-\tau)}}\cdot\mathbb{I}\{\tau<T\}\right]\\
=&\mathbb{E}_t\left[V(W_\tau-R_1,\ \tau)\cdot e^{-\gamma\tilde{R}_\tau e^{r(T-\tau)}}\cdot\mathbb{I}\{\tau<T\}\right]\\
=&\mathbb{E}_t\left[V(W_\tau,\tau)\cdot e^{\gamma R_1 e^{r(T-\tau)}}\cdot e^{-\gamma\tilde{R}_\tau e^{r(T-\tau)}}\cdot\mathbb{I}\{\tau<T\}\right]\\
=&\mathbb{E}_t\left[V(W_\tau,\tau)\cdot\mathbb{E}_\tau\ e^{\gamma R_1 e^{r(T-\tau)}}\cdot e^{-\gamma\tilde{R}_\tau e^{r(T-\tau)}}\cdot\mathbb{I}\{\tau<T\}\right]\\
=&\mathbb{E}_t\left[V(W_\tau,\tau)\cdot\mathbb{I}\{\tau<T\}\right]\\
\geq&\mathbb{E}_t\left[u(W_T)\cdot\mathbb{I}\{\tau<T\}\right].
\end{aligned}
$$

Consequently, we get $H(w,C,m,t)\geq\mathbb{E}_t[W_T]$ for all admissible trading strategies, and hence $H\geq U$. Again we get equality for $\boldsymbol{\pi}=\overline{\boldsymbol{\pi}}$ which yields $H=U$. $\square$

Chapter 5

Indifference Pricing under Model Misspecification

5.1 Motivation

Following ideas of Anderson, Hansen, and Sargent (2000) and Maenhout (2004), we now incorporate model uncertainty into the default model from chapter 3. Suppose that in a market with a money market account and n risky, tradable and default free assets $S_t^{(1)}, \ldots, S_t^{(n)}$ the investor seeks to maximize expected utility of terminal wealth over all admissible trading strategies, i.e. wants to determine $\sup_{\boldsymbol{\pi} \in \mathcal{A}} \mathbb{E}[u(W_T)]$. The dynamics of the economy is first estimated to be described by the measure $\mathbb{P}$, called the *reference measure*, under which $S_t^{(1)}, \ldots, S_t^{(n)}$ have the dynamics

$$dS_t^{(i)} = S_t^{(i)} \left(\mu_t^{(i)} dt + \sigma_t^{(i)} dB_t^{(i)} \right),$$

the $B_t^{(i)}$ are correlated Wiener processes. Since the investor is uncertain whether or not $\mathbb{P}$ is indeed the correct measure, he is willing to consider other *candidate measures* $Q \sim \mathbb{P}$ as well. However, a measure change comes at the cost of a penalty for his value function,

and his new goal is to find

$$\sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \left\{\mathbb{E}^Q[u(W_T)] + h(Q)\right\} \tag{5.1}$$

In equation (5.1) the freedom of choice of the candidate measure can be interpreted as a second control (apart from the trading strategy). The penalty term $h(Q)$ controls the distance between the candidate measure Q and the reference measure $\mathbb{P}$ which the investor still considers reasonable.

A popular choice for h (see e.g. Anderson, Hansen, and Sargent (2000)) is the entropic penalty function

$$h(Q) = k\ \mathcal{H}(Q|\mathbb{P}) = k\ \mathbb{E}^Q\left[\log\frac{dQ}{d\mathbb{P}}\right] = k\ \mathbb{E}\left[\frac{dQ}{d\mathbb{P}}\log\frac{dQ}{d\mathbb{P}}\right],$$

where k is a positive constant.

If we work with the standard filtration generated by the driving Brownian motions $B_t^{(1)},\ \ldots,\ B_t^{(n)}$, augmented by the sets of measure 0 of Ω, then an equivalent measure change corresponds to an adjustment of the drifts in the dynamics of the $S_t^{(i)}$, i.e. under a measure $Q\sim\mathbb{P}$ the risky assets have the dynamics

$$dS_t^{(i)} = S_t^{(i)}\left[(\mu_t^{(i)} + v_t^{(i)})\ dt + \sigma_t^{(i)}\ dB_t^{Q,(i)}\right] \tag{5.2}$$

for some $\mathcal{F}_t$-adapted processes $v_t^{(1)},\ \ldots,\ v_t^{(n)}$. Here the $B_t^{Q,(i)}$ are correlated Brownian motions under Q.

A short calculation shows that

$$\mathbb{E}^Q\left[\log\frac{dQ}{d\mathbb{P}}\right] = \frac{1}{2}\ \mathbb{E}^Q\left[\ \int_0^T \boldsymbol{v}_s^T\boldsymbol{\Omega}^{-1}\boldsymbol{v}_s\ ds\ \right],$$

where $\boldsymbol{v}_t = (v_t^{(1)},\ldots,v_t^{(n)})^T$ and $\boldsymbol{\Omega}$ is the variance-covariance matrix of $S_t^{(1)},\ \ldots,\ S_t^{(n)}$.

Let us restrict ourselves to candidate measures corresponding to Markovian drift adjustments, i.e.

$$\boldsymbol{v}_t = \boldsymbol{v}(\boldsymbol{S}_t, t), \quad \boldsymbol{S}_t = \left(S_t^{(1)},\ \ldots,\ S_t^{(n)}\right).$$

It follows, that if U is the value function defined by (5.1), then U has the HJB equation

$$\partial_t U + \sup_{\boldsymbol{\pi}\in\mathbb{R}^n} \inf_{\boldsymbol{v}\in\mathbb{R}^n} \left\{ \mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}} U + \frac{1}{2} k \; \boldsymbol{v}^T \boldsymbol{\Omega}^{-1} \boldsymbol{v} \right\} = 0, \tag{5.3}$$

subject to the appropriate terminal condition $U(w, \boldsymbol{S}, T) = u(w)$. Here $\mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}}$ is the infinitesimal generator of the joint process $(W_t, \boldsymbol{S}_t)$.

Unfortunately, due to the last term on the left hand side not containing U or any of its derivatives, equation (5.3) cannot be solved analytically. In fact, for the case of power or exponential utility, we are even unable to factor wealth out of the solution. As a resolution, Maenhout (2004) suggests the following approach: he modifies the HJB equation by scaling the penalty term and defines U to be the solution of the equation

$$\partial_t U + \sup_{\boldsymbol{\pi}\in\mathbb{R}^n} \inf_{\boldsymbol{v}\in\mathbb{R}^n} \left\{ \mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}} U + \frac{1}{2} k U \; \boldsymbol{v}^T \boldsymbol{\Omega}^{-1} \boldsymbol{v} \right\} = 0. \tag{5.4}$$

Maenhout then shows that this modified equation can be solved and uses it to optimize portfolios with an infinite time horizon and consumption-based power utility.

Alternatively one can also modify the optimization problem and define the value function as the solution of

$$U(w,t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2} k \int_t^T U(W_s, s) \; \boldsymbol{v}_s^T \boldsymbol{\Omega}^{-1} \boldsymbol{v}_s \; ds \;\middle|\; W_t = w \right]. \tag{5.5}$$

If we define U in this way, existence and uniqueness questions are obviously raised. These issues are discussed in Duffie and Epstein (1992) in the context of differential utility.

The penalty term equation (5.4) can be interpreted as a scaled version of the penalty term corresponding to relative entropy. However the scaling factor is not constant, but dependent on future utility.

The definition of U would normally also depend on initial conditions corresponding to the risky assets. However, as in the case with complete certainty, it turns out that e.g. for power and exponential utiliy, U depends only on wealth and time. The HJB-equation corresponding to the modification (5.5) is in fact equation (5.4).

In the following sections we adopt this approach to generalize our results from the previous sections on the investment problem as well as on pricing the defaultable bond and the credit default swap. Our work differs from Maenhout (2004) firstly in the choice of the utility function. More importantly, due to the default risk of one of the assets, the arising optimization problems are different and lead to more complicated HJB equations. Nevertheless, we show that we can solve them analytically, at least to the same extent as for the case of complete model specification.

In a related paper, Uppal and Wang (2003) generalize the setting in Maenhout (2004) and introduce different levels of ambiguity for different assets. Applying their idea to our setting, we define the value function U as solution of the equation

$$U(w,t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2}\int_t^T U(W_s,s)\ \boldsymbol{v}_s^T \boldsymbol{\Phi} \boldsymbol{v}_s\ ds \ \middle|\ W_t = w \right]. \tag{5.6}$$

where the matrix $\boldsymbol{\Phi}$ arises as a weighted sum of the levels of model uncertainty corresponding to different subsets of the risky assets. The construction of $\boldsymbol{\Phi}$ is discussed in detail in appendix 5.B. This generalization is particularly useful for our setting, since the model for the tradable assets can usually be estimated well from past data, whereas the dynamics of the health of the defaultable firm are rather uncertain.

The default model we use in this section is the same as the one from chapter 3. In particular, there are three possible states for the firm (healthy, distressed and default). The switch from *healthy* to *distressed* is triggered by the creditworthiness index C hitting the barrier D for the first time, while default is triggered by the switch of a Poisson process. The only difference compared to chapter 3 is that we assume uncertainty on the dynamics of the tradable risky assets and the CWI.

5.2 The Investment Problem

For the remainder of this chapter we assume that under a measure $Q \sim \mathbb{P}$ the dynamics of I, S, C are given by

$$\begin{cases} dI_t = I_t \left[(\mu_1 + v_t^{(1)})\ dt + \sigma_1\ dB_t^{Q,(1)} \right] , \\ dS_t = S_t \left[(\mu_2 + v_t^{(2)})\ dt + \sigma_2\ dB_t^{Q,(2)} \right] , \\ dC_t = C_t \left[(\nu + v_t^{(3)})\ dt + \eta\ dB_t^{Q,(3)} \right] \end{cases} \tag{5.7}$$

with correlated Q-Brownian motions $B_t^{Q,(1)}$, $B_t^{Q,(2)}$, $B_t^{Q,(3)}$. Normally we will write $B_t^{(1)}$, $B_t^{(2)}$, $B_t^{(3)}$ only for notational purposes. In addition to the same notation as in section 3.3 we let $\boldsymbol{v}_t = (v_t^{(1)}, v_t^{(2)}, v_t^{(3)})^T$. As in the completely specified case, we assume that the investor liquidates their position in S at time τ_h. Consequently, in the absence of default risk, the corresponding wealth process has the Q-dynamics

$$dW_t = \begin{cases} \left[((\boldsymbol{\mu} - \boldsymbol{r})^T + (v_t^{(1)}, v_t^{(2)}))\boldsymbol{\pi}_t + r\ W_t \right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t < \tau_h, \\ \left[(\mu_1 - r + v_t^{(1)})\pi_t^{(1)} + r\ W_t \right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & t > \tau_h, \end{cases} \tag{5.8}$$

subject to

$$W_{\tau_h} = W_{\tau_h^-}.$$

We first define the set of admissible trading strategies, denoted again by $\mathcal{A}$, as well as the set of *candidate measures*. We could define $\mathcal{A}$ similarly to section 3. To keep the definition simple, here we require from the start that each trading strategy and every drift adjustment corresponding to a candidate measure be bounded.

Definition 6. *(a) An admissible trading strategy is a predictable stochastic process* $\boldsymbol{\pi}_t = \left(\pi_t^{(1)}, \pi_t^{(2)}\right) \in \mathbb{R}^2$ *which is almost surely bounded and satisfies* $\pi_t^{(2)} = 0$ *for* $t > \tau_h$.

(b) A candidate measure is a measure $Q \sim \mathbb{P}$ *under which* I, S, C *have the dynamics (5.7), where* $\boldsymbol{v}_t = (v_t^{(1)}, v_t^{(2)}, v_t^{(3)})$ *is almost surely bounded and predictable.*

5.2.1 The Value Function in the Distressed Regime

We start with computing the value function V corresponding to an optimal investment in the distressed regime. For $t > \tau_h$ the wealth process has the Q-dynamics

$$dW_t = \left[rW_t + (\mu_1 - r + v_t^{(1)})\pi_t^{(1)}\right]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}\ .$$

From this point forward, we will write v and π instead of $v^{(1)}$ and $\pi^{(1)}$, when there is no confusion.

We define V to be the solution of the optimization problem

$$V(w, I, t) = \sup_{\pi\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2}\int_t^T V(W_s, I_s, s)\ \phi\ v_s^2\ ds \ \middle|\ W_s = w,\ I_s = I\right]. \quad (5.9)$$

Here ϕ is a *negative scalar*. The penalty term in (5.9) is a scaled version of the relative entropy of a measure change induced by adjusting the drift of the index I only. We choose this penalty, because in the distressed regime the only model uncertainty relevant for the wealth process is through $v_t^{(1)}$, the drift adjustment of I. The scalar ϕ is negative, because V is negative.

Assuming V is independent of I, the corresponding HJB equation is

$$\begin{cases} \partial_t V + \sup_{\pi\in\mathbb{R}} \inf_{v\in\mathbb{R}} \left\{\mathcal{L}^{\pi,v}\ V + \frac{1}{2}\ V\phi v^2\right\} = 0\ , \\ \qquad\qquad\qquad\qquad V(w,T) = u(w)\ , \quad w\in\mathbb{R}\ , \end{cases} \quad (5.10)$$

where

$$\mathcal{L}^{\pi,v}V = (rw + (\mu_1 - r + v)\pi)\ \partial_w V + \frac{1}{2}\pi^2\sigma_1^2\ \partial_{ww}V.$$

In contrast to the completely specified case, it is by far not obvious that V is indeed independent of I. However the verification theorem from appendix 5.A shows that the solution of equation (5.10) that we will find indeed coincides with the value function V in equation (5.9).

As in the completely specified case, we make the ansatz $V(w,t) = u(we^{r(T-t)})\ g(t)$,

which results in the following ODE for g:

$$\begin{cases} g' + \inf_{\pi\in\mathbb{R}} \sup_{v\in\mathbb{R}} \left\{ g(\mu_1 - r + v)\ a_t\pi + \frac{1}{2}\pi^2\sigma_1^2 a_t^2\ g + \frac{1}{2} g\ \phi\ v^2 \right\} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g(T) = 1. \end{cases} \tag{5.11}$$

To find the saddle point, it is convenient to write the PDE in the form

$$g' + \inf_{\pi\in\mathbb{R}} \sup_{v\in\mathbb{R}} F(\boldsymbol{h}) = 0, \tag{5.12}$$

where

$$F(\boldsymbol{h}) = \frac{1}{2}\boldsymbol{h}^T\boldsymbol{K}\boldsymbol{h} + \boldsymbol{d}^T\boldsymbol{h}$$

and

$$\boldsymbol{h} = \begin{pmatrix} v \\ \tilde{\pi} \end{pmatrix}, \quad \tilde{\pi} = a_t\pi, \quad \boldsymbol{K} = g\begin{pmatrix} \phi & 1 \\ 1 & \sigma_1^2 \end{pmatrix}, \quad \boldsymbol{d} = \begin{pmatrix} 0 \\ (\mu_1 - r)g \end{pmatrix}.$$

The first order condition for a saddle point is $\boldsymbol{K}\boldsymbol{h} = -\boldsymbol{d}$ which has the unique solution

$$\boldsymbol{h}^* = -\boldsymbol{K}^{-1}\boldsymbol{d} = \frac{\mu_1 - r}{\phi\sigma_1^2 - 1}\begin{pmatrix} -1 \\ \phi \end{pmatrix}.$$

Since $\phi < 0$, the determinant of $\boldsymbol{K}$ is negative and therefore $\boldsymbol{h}^*$ indeed corresponds to a saddle point.

The corresponding critical value is $F(\boldsymbol{h}^*) = -\frac{1}{2}\boldsymbol{d}^T\boldsymbol{K}^{-1}\boldsymbol{d}$, and leads to the ODE

$$\begin{cases} g' - \frac{1}{2}\frac{(\mu_1 - r)^2\phi}{\phi\sigma_1^2 - 1}\ g = 0, \\ \qquad\qquad\qquad g(T) = 1. \end{cases} \tag{5.13}$$

This ODE has the solution $g(t) = e^{-\frac{1}{2}\overline{\lambda}^2(T-t)}$ with $\overline{\lambda}^2 = \frac{(\mu_1 - r)\phi}{(\phi\sigma_1^2 - 1)}$.

The above result is quite interesting in the two extreme cases $\phi = 0$ and $\phi \to -\infty$. In the limit $\phi \to -\infty$ the solution reduces to the standard Merton problem with complete certainty. This is expected, since $\phi \to -\infty$ penalizes any deviation from the reference measure $\mathbb{P}$ heavily. In contrast, $\phi = 0$ corresponds a complete lack of confidence in the reference measure $\mathbb{P}$. It is interesting to observe that in this case, $\pi^* \to 0$, i.e. the

investor invests less and less money in the risky asset. Furthermore, $v_t = r - \mu_1$, and the corresponding measure Q is measure under which the risky asset grows at the risk-free rate.

5.2.2 The Value Function in the Healthy Regime

Prior to τ_h the corresponding wealth process has the Q-dynamics

$$dW_t = \left[\left((\boldsymbol{\mu} - \boldsymbol{r})^T + (v_t^{(1)}, v_t^{(2)})\right)\boldsymbol{\pi}_t + r\ W_t\right]\ dt + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)},$$

The value function $U(w, I, S, C, t)$ in the healthy regime is defined to be the solution of the optimization problem

$$\begin{aligned} U(w, I, S, C, t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \Bigg[u(W_T) + \frac{1}{2}\int_t^{\tau_h\wedge T} U(W_s, P_s, S_s, C_s, s)\ \boldsymbol{v}_s^T \boldsymbol{\Phi} \boldsymbol{v}_s\ ds\ + \\ + \frac{1}{2}\int_{\tau_h\wedge T}^T V(W_s, s)\ \phi\ v_s^2\ ds \ \Bigg|\ W_t = w,\ I_t = I,\ S_t = S,\ C_t = C \Bigg]. \quad (5.14) \end{aligned}$$

Here $\boldsymbol{\Phi}$ is a *negative semidefinite* matrix. Its construction is explained in detail in appendix 5.B.

The choice of the penalty terms in equation (5.14) are motivated in a similar way as in equation (5.9). Up to time τ_h, the penalty is a scaled version of the relative entropy of the measure change from $\mathbb{P}$ to Q, since the wealth process is affected by model uncertainty in all three processes I_t, S_t, C_t. After time τ_h only model uncertainty in I affects W_t, so we choose the same penalty term as in the value function for the distressed regime.

Assuming that U is dependent on w, C, and t only, the corresponding HJB equation is

$$\begin{cases} \partial_t U + \sup\limits_{\boldsymbol{\pi}\in\mathbb{R}^2} \inf\limits_{\boldsymbol{v}\in\mathbb{R}^3} \mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}} U = 0, \\ U(w, C, T) = u(w) \\ U(w, D, t) = V(w, t), \end{cases} \quad (5.15)$$

with

$$\begin{aligned}\mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}}U = &\left[rw + \boldsymbol{\pi}^T\left((\boldsymbol{\mu}-\boldsymbol{r}) + (v^{(1)}, v^{(2)})^T\right)\right]\partial_w\, U + \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\;\partial_{ww}U + \\ &+ (\nu + v^{(3)})C\;\partial_C U + \boldsymbol{\pi}^T\boldsymbol{\omega}C\;\partial_{wC} + \frac{1}{2}\eta^2C^2\;\partial_{CC}U + \frac{1}{2}\;U\;\boldsymbol{v}^T\boldsymbol{\Phi}\boldsymbol{v}\end{aligned} \tag{5.16}$$

Once again, it is not obvious that U should depend only on w, C and t; however, the verification theorem from appendix 5.A demonstrates that the solution to (5.15) indeed coincides with the value function U in (5.14).

The method for solving this equation is similar to the one in section 3.3. Our goal is to write U in the form $U(w,C,t) = u(we^{r(T-t)})\;e^{-\frac{1}{2}\overline{\Lambda}^2(T-t)}\;G(\ln\frac{C}{D}, T-t)^\beta$. Then we determine $\overline{\Lambda}$ and β such that the PDE for G is linear. Firstly, letting $U(w,C,t) = u(we^{r(T-t)})\;g(C,t)$ leads to the PDE

$$\begin{cases}\partial_t g + \inf\limits_{\boldsymbol{\pi}\in\mathbb{R}^2}\;\sup\limits_{\boldsymbol{v}\in\mathbb{R}^3}\left\{\boldsymbol{\pi}^T\left(\boldsymbol{\mu}-\boldsymbol{r}+(v^{(1)},v^{(2)})^T\right)a_t g + \frac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\;a_t^2 g + \right. \\ \left. +(\nu+v^{(3)})C\;\partial_C g + \boldsymbol{\pi}^T\boldsymbol{\omega}\;a_t C\;\partial_C g + \frac{1}{2}\eta^2C^2\;\partial_{CC}g + \frac{1}{2}g\;\boldsymbol{v}^T\boldsymbol{\Phi}\boldsymbol{v}\right\} = 0, \\ g(C,T) = 1 \\ g(D,t) = e^{-\frac{1}{2}\overline{\lambda}^2(T-t)}.\end{cases} \tag{5.17}$$

As in the case of the distressed regime, the notation for this equation can be simplified by writing it in the form

$$\partial_t g + \nu\;C\partial_C g + \frac{1}{2}\eta^2C^2\;\partial_{CC}g + \inf_{\boldsymbol{\pi}\in\mathbb{R}^2}\;\sup_{\boldsymbol{v}\in\mathbb{R}^3}\;F(\boldsymbol{h}) = 0,$$

where

$$F(\boldsymbol{h}) = \frac{1}{2}\;\boldsymbol{h}^T\boldsymbol{K}\boldsymbol{h} + \boldsymbol{d}^T\boldsymbol{h}$$

and

$$\boldsymbol{h} = \begin{pmatrix}\boldsymbol{v}\\ \tilde{\boldsymbol{\pi}}\end{pmatrix}, \quad \tilde{\boldsymbol{\pi}} = a_t\boldsymbol{\pi}, \quad \boldsymbol{K} = g\begin{pmatrix}\boldsymbol{\Phi} & \begin{matrix}1 & 0\\ 0 & 1\\ 0 & 0\end{matrix} \\ \begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\end{matrix} & \boldsymbol{\Omega}\end{pmatrix}, \quad \boldsymbol{d} = \begin{pmatrix}0\\ 0\\ C\;\partial_C g\\ (\boldsymbol{\mu}-\boldsymbol{r})g + \boldsymbol{\omega}C\;\partial_C g\end{pmatrix}.$$

As before, the optimization problem has the unique solution $\boldsymbol{h}^* = -\boldsymbol{K}^{-1}\boldsymbol{d}$, the corresponding critical value is $F(\boldsymbol{h}^*) = -\frac{1}{2}\,\boldsymbol{d}^T\boldsymbol{K}^{-1}\boldsymbol{d}$, and the PDE reduces to

$$\partial_t g + \nu C\partial_C g + \frac{1}{2}\,\eta^2 C^2 \partial_{CC} g - \frac{1}{2}\,\boldsymbol{d}^T\boldsymbol{K}^{-1}\boldsymbol{d} = 0. \tag{5.18}$$

To solve this equation, we first compute $\boldsymbol{K}^{-1}$. Letting $\boldsymbol{E} \triangleq \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ and noting that $\boldsymbol{\Phi}$ and $\boldsymbol{\Omega}$ are symmetric, we have

$$\boldsymbol{K}^{-1} = \frac{1}{g}\left(\begin{array}{c|c} (\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T)^{-1} & -\boldsymbol{\Phi}^{-1}\boldsymbol{E}(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1} \\ \hline -(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1}\boldsymbol{E}^T\boldsymbol{\Phi}^{-1} & (-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1} \end{array}\right) \tag{5.19}$$

Since the term $-\frac{1}{2}\boldsymbol{d}^T\boldsymbol{K}^{-1}\boldsymbol{d}$ contributes the term $-\frac{1}{2}g\,(\boldsymbol{\mu}-\boldsymbol{r})^T(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E}+\boldsymbol{\Omega})^{-1}(\boldsymbol{\mu}-\boldsymbol{r})$ to equation (6.5), we make the substitution

$$g(C,t) = G^{\beta}(\ln\frac{C}{D}, T-t))\; e^{-\frac{1}{2}\overline{\Lambda}^2(T-t)},$$

where

$$\overline{\Lambda}^2 = (\boldsymbol{\mu}-\boldsymbol{r})^T(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1}(\boldsymbol{\mu}-\boldsymbol{r}).$$

To find the resulting equation for G it is convenient to write

$$\boldsymbol{d} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \boldsymbol{\mu}-\boldsymbol{r} \end{pmatrix} g + \begin{pmatrix} 0 \\ 0 \\ 1 \\ \boldsymbol{\omega} \end{pmatrix} C\partial_C g,$$

Using the notation $\boldsymbol{e} = (0,0,1)^T$, G satisfies the PDE

$$\begin{aligned} -\partial_\tau G + \left(\nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu}-\boldsymbol{r})^T\left(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega}\right)^{-1}\left(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{e} + \boldsymbol{\omega}\right)\right)\;\partial_x G + \frac{1}{2}\eta^2\;\partial_{xx}G - \\ + \frac{1}{2}\frac{(\partial_x G)^2}{G}\left[(\beta-1)\eta^2 - \beta\;\widetilde{\boldsymbol{d}}^T\boldsymbol{K}^{-1}\widetilde{\boldsymbol{d}}\right] \;=\; 0, \end{aligned}$$

where $\widetilde{d} = \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{\omega} \end{pmatrix}$, and hence the appropriate choice of β is

$$\beta = \frac{1}{1 - \frac{1}{\eta^2}\, \widetilde{\boldsymbol{d}}^T \boldsymbol{K}^{-1} \widetilde{\boldsymbol{d}}}.$$

With this choice of β the PDE for G is linear:

$$\begin{cases} -\partial_\tau G + \widetilde{\nu}\, \partial_x G + \dfrac{1}{2}\eta^2\, \partial_{xx} G = 0, \\ \qquad\qquad G(0,\tau) = e^{-\frac{1}{2\beta}(\overline{\Lambda}^2 - \overline{\lambda}^2)\tau}, \\ \qquad\qquad G(x,0) = 1, \end{cases} \tag{5.20}$$

where this time $\widetilde{\nu} = \nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu} - \boldsymbol{r})^T \left(-E^T \boldsymbol{\Phi}^{-1} E + \boldsymbol{\Omega}\right)^{-1} \left(-\boldsymbol{E}^T \boldsymbol{\Phi}^{-1} \boldsymbol{e} + \boldsymbol{\omega}\right)$. It is pleasing that the same PDE that arises in the case of full certainty (equation (3.3), section 3.3) appears here as well – except with λ, Λ replaced by $\overline{\lambda}$, $\overline{\Lambda}$ and with a modified $\tilde{\nu}$ (no longer the MEMM adjusted drift).

As for the distressed regime, we consider the limiting behaviour of the value function for the cases $\boldsymbol{\Phi} \to -\infty$ and $\boldsymbol{\Phi} \to 0$. For convenience we restate the values of the parameters in equation (5.20):

$$\begin{aligned} \widetilde{\nu} &= \nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu} - \boldsymbol{r})^T \left(-E^T \boldsymbol{\Phi}^{-1} E + \boldsymbol{\Omega}\right)^{-1} \left(-\boldsymbol{E}^T \boldsymbol{\Phi}^{-1} \boldsymbol{e} + \boldsymbol{\omega}\right), \\ \beta &= \frac{1}{1 - \frac{1}{\eta^2}\, \widetilde{d}^T \boldsymbol{K}^{-1} \widetilde{d}}, \\ \overline{\Lambda}^2 &= (\boldsymbol{\mu} - \boldsymbol{r})^T (-\boldsymbol{E}^T \boldsymbol{\Phi}^{-1} \boldsymbol{E} + \boldsymbol{\Omega})^{-1} (\boldsymbol{\mu} - \boldsymbol{r}), \\ \overline{\lambda}^2 &= \frac{(\mu_1 - r)^2 \phi}{\phi \sigma_1^2 - 1}. \end{aligned}$$

For $\boldsymbol{\Phi} \to -\infty$, $\phi \to -\infty$ it is easy to see that these parameters converge to the corresponding parameters in the completely specified case, and hence the same applies for the value function.

For the other extreme case we first explain what we mean by $\boldsymbol{\Phi} \to 0$. Let $\boldsymbol{\Phi} = \varepsilon \boldsymbol{\Phi}_0$, $\phi = \varepsilon \phi_0$, for some fixed $\boldsymbol{\Phi}_0$ and ϕ_0. We will examine the behaviour of the value function

as $\varepsilon \to 0$. It is convenient to write $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Phi}_0^{-1}$ in the form

$$\boldsymbol{\Phi}_0 = \left(\begin{array}{c|c} \boldsymbol{\Psi}_0 & \boldsymbol{a} \\ \hline \boldsymbol{a}^T & \varphi_0 \end{array}\right), \quad \boldsymbol{\Phi}_0^{-1} = \left(\begin{array}{c|c} \overline{\boldsymbol{\Psi}}_0^{-1} & \overline{\boldsymbol{a}} \\ \hline \overline{\boldsymbol{a}}^T & \overline{\varphi}_0 \end{array}\right).$$

If we assume that $\boldsymbol{\Phi}_0$ is strictly negative definite, then $\varphi_0 < 0$, $\overline{\varphi}_0 < 0$ and $\boldsymbol{\Psi}_0$, $\overline{\boldsymbol{\Psi}}_0^{-1}$ are strictly negative definite.

Since the entries of $\boldsymbol{K}^{-1}$ frequently appear in the parameters above, we first examine their behaviour for $\varepsilon \to 0$. We immediately see that

$$(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E}+\boldsymbol{\Omega})^{-1} = (-\varepsilon\overline{\boldsymbol{\Psi}}_0^{-1}+\boldsymbol{\Omega})^{-1} \;=\; -\varepsilon\overline{\boldsymbol{\Psi}}_0(I-\varepsilon\boldsymbol{\Omega}\overline{\boldsymbol{\Psi}}_0)^{-1} \;=\; -\varepsilon\overline{\boldsymbol{\Psi}}_0+O(\varepsilon^2), \quad (5.21)$$

and

$$\boldsymbol{\Phi}^{-1}\boldsymbol{E}(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E}+\boldsymbol{\Omega}^{-1})^{-1} = \frac{1}{\varepsilon}\left(\begin{array}{c} \overline{\boldsymbol{\Psi}}_0^{-1} \\ \hline \overline{\boldsymbol{a}}^T \end{array}\right)\cdot\left[-\varepsilon\overline{\boldsymbol{\Psi}}_0 + O(\varepsilon^2)\right] = \left(\begin{array}{c} -\boldsymbol{I} \\ \hline -\overline{\boldsymbol{a}}^T\overline{\boldsymbol{\Psi}}_0 \end{array}\right)+O(\varepsilon). \quad (5.22)$$

Now we examine the behaviour of the entries of $(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}\boldsymbol{E}^T)^{-1}$.

Lemma 15. *The (3,3) entry of* $(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}\boldsymbol{E}^T)^{-1}$ *is*

$$\frac{1}{\varepsilon\varphi_0} + O(1), \quad (5.23)$$

and all other entries of $(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T)^{-1}$ *approach finite values as* $\varepsilon \to 0$.

Proof. Recall that for a regular square matrix $\boldsymbol{A} = (a_{ij})$, the elements of the inverse matrix $\boldsymbol{A}^{-1} = (\overline{a}_{ij})$ can be computed as follows:

$$\overline{a}_{ij} = \frac{(-1)^{i+j}\cdot[(j,i)\text{ minor of }\boldsymbol{A}]}{\det\boldsymbol{A}}.$$

By (i,j) minor of $\boldsymbol{A}$ we mean the determinant of the submatrix obtained from $\boldsymbol{A}$ by deleting the ith row and the jth column.

We have

$$\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}\boldsymbol{E}^T = \left(\begin{array}{c|c} \varepsilon\boldsymbol{\Psi}_0 - \boldsymbol{\Omega}^{-1} & \varepsilon\boldsymbol{a} \\ \hline \varepsilon\boldsymbol{a}^T & \varepsilon\varphi_0 \end{array}\right),$$

and hence

$$\det(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}\boldsymbol{E}^T) = \varepsilon\varphi_0 \cdot \det \boldsymbol{\Omega}^{-1} + O(\varepsilon^2).$$

The (3,3) minor of $\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T$ is

$$\det(\varepsilon\boldsymbol{\Psi}_0 - \boldsymbol{\Omega}^{-1}) = \det \boldsymbol{\Omega}^{-1} + O(\varepsilon),$$

and therefore the (3,3) entry of $(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}\boldsymbol{E}^T)^{-1}$ is

$$\frac{1}{\varepsilon\varphi_0} + O(1). \tag{5.24}$$

All other minors of $\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T$ are at least of order $O(\varepsilon)$, and therefore all other entries of $(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T)^{-1}$ approach finite values as $\varepsilon \to 0$. □

From the results of (5.21), (5.22) it follows that as $\varepsilon \to 0$,

$$\begin{aligned}\widetilde{\nu} &= \nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu} - \boldsymbol{r})^T \left(-\varepsilon\overline{\boldsymbol{\Psi}}_0 + O(\varepsilon^2)\right)\left(\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{e} + \boldsymbol{\omega}\right) \\ &= \nu - \frac{1}{2}\eta^2 - (\boldsymbol{\mu} - \boldsymbol{r})^T \left(-\varepsilon\overline{\boldsymbol{\Psi}}_0 + O(\varepsilon^2)\right)(\varepsilon^{-1}\overline{\boldsymbol{a}} + \boldsymbol{\omega}) \quad \to \quad \nu - \frac{1}{2}\eta^2 + (\boldsymbol{\mu} - \boldsymbol{r})^T\overline{\boldsymbol{\Psi}}_0\overline{\boldsymbol{a}}\end{aligned}$$

and

$$\begin{aligned}\overline{\Lambda}^2 &= -\varepsilon\ (\boldsymbol{\mu} - \boldsymbol{r})^T\overline{\boldsymbol{\Psi}}_0(\boldsymbol{\mu} - \boldsymbol{r}) + O(\varepsilon^2), \\ \overline{\lambda}^2 &= -\varepsilon\ (\mu_1 - r)^2\ \phi_0 + O(\varepsilon^2).\end{aligned}$$

Furthermore, the only entry of $\boldsymbol{K}^{-1}$ of order $O(\varepsilon^{-1})$ is the $(3,3)$ entry. Hence $\beta \to 0$ ($\varepsilon \to 0$), and more precisely by lemma 15, we have

$$\beta = -\varepsilon\eta^2\varphi_0 + O(\varepsilon^2).$$

It follows that

$$\frac{1}{\beta}(\overline{\Lambda}^2 - \overline{\lambda}^2) \to \frac{1}{\eta^2\varphi_0} \cdot \left[(\boldsymbol{\mu} - \boldsymbol{r})^T\overline{\boldsymbol{\Psi}}_0(\boldsymbol{\mu} - \boldsymbol{r}) - (\mu_1 - r)^2\ \phi_0\right].$$

Therefore the boundary conditions in equation (5.20) imply that G remains bounded as well as bounded away from 0 (as $\varepsilon \to 0$). Since $\beta \to 0$, it follows that $g \to 1$ and hence $U \to u(we^{r(T-t)})$.

Now we examine how the optimal trading strategy $\boldsymbol{\pi}^*$ and the optimal measure Q^* behave as $\varepsilon \to 0$. Recall that $\begin{pmatrix} \boldsymbol{v}^* \\ \boldsymbol{\pi}^* \end{pmatrix} = -\boldsymbol{K}^{-1}\boldsymbol{d}$. Firstly, it is helpful to notice that

$$\frac{C\partial_C g}{g} = \beta \, \frac{\partial_x G}{G} \to 0 \quad (\varepsilon \to 0),$$

since in the limit as $\epsilon \to 0$: $\partial_x G$ is bounded, G is bounded away from 0 and $\beta \to 0$.

Theorem 16. *For $\varepsilon \to 0$,*

$$\pi_t^{(1),*} \to 0, \;\; \pi_t^{(2),*} \to 0$$

pointwise for all w, I, S, C, t.

In other words, when there is complete uncertainty, no risky investments are made. This is the analog to the distressed regime's result.

Proof. For $\boldsymbol{\pi}^*$ we get from equation (5.19)

$$\boldsymbol{\pi}^* = (-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1}(\boldsymbol{\mu} - \boldsymbol{r}) - (-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1}\left(\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{e} + \boldsymbol{\omega}\right)\frac{C\partial_C g}{g}$$

and hence

$$\boldsymbol{\pi}^* = -\varepsilon\overline{\boldsymbol{\Psi}}_0(\boldsymbol{\mu} - \boldsymbol{r}) + \varepsilon\overline{\boldsymbol{\Psi}}_0(\varepsilon^{-1}\overline{\boldsymbol{a}} + \boldsymbol{\omega})\frac{C\;\partial_C g}{g} + O(\varepsilon) \quad \to \quad \mathbf{0}.$$

□

For the optimal measure we first focus on the limiting behaviour of $v_t^{(1)}$ and $v_t^{(2)}$.

Theorem 17. *As $\varepsilon \to 0$,*

$$v_t^{(1),*} \to r - \mu_1, \quad v_t^{(2),*} \to r - \mu_2$$

pointwise for all w, I, S, C, t.

Hence under the optimal measure, the drifts of P_t and S_t tend to r as uncertainty increases. This result is also the analog to the distressed regime.

Proof. For $\boldsymbol{v}^*$ we get from equation (5.19)

$$\begin{aligned}\boldsymbol{v}^* &= -(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T)^{-1}\boldsymbol{e}\frac{C\partial_C g}{g} + \boldsymbol{\Phi}^{-1}\boldsymbol{E}(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1}\ (\boldsymbol{\mu} - \boldsymbol{r}) + \\ &\quad + \boldsymbol{\Phi}^{-1}\boldsymbol{E}(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1}\ \boldsymbol{\omega}\frac{C\partial_C g}{g}. \qquad (5.25)\end{aligned}$$

Since $\boldsymbol{\Phi}^{-1} = \varepsilon^{-1}\boldsymbol{\Phi}_0^{-1}$, $(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega})^{-1} = -\varepsilon\overline{\boldsymbol{\Psi}}_0 + O(\varepsilon^2)$, $\frac{C\partial_C g}{g} \to 0$, the third term approaches 0 as $\varepsilon \to 0$. Furthermore, since by (5.22),

$$\boldsymbol{\Phi}^{-1}\boldsymbol{E}(-\boldsymbol{E}^T\boldsymbol{\Phi}^{-1}\boldsymbol{E} + \boldsymbol{\Omega}^{-1})^{-1} = \begin{pmatrix} -\boldsymbol{I} \\ \hline -\overline{\boldsymbol{a}}^T\overline{\boldsymbol{\Psi}}_0 \end{pmatrix} + O(\varepsilon),$$

it is easy to see that the first two components of the second term tend to $r - \mu_1$ and $r - \mu_2$. Therefore we still have to show that the first two components of the first term approach 0. Recalling that $\boldsymbol{e} = (0, 0, 1)^T$, this follows from the fact that the (1,3) and (2,3) entries of $(\boldsymbol{\Phi} - \boldsymbol{E}\boldsymbol{\Omega}^{-1}\boldsymbol{E}^T)^{-1}$ approach finite values for $\varepsilon \to 0$. □

Theorem 18. *Let $L(x,t) \triangleq \frac{\partial_x G}{G}$. Then for $\varepsilon \to 0$,*

$$v_t^{(3),*} \to \eta^2 L(x,t) - \overline{\boldsymbol{a}}^T\overline{\boldsymbol{\Psi}}_0(\boldsymbol{\mu} - \boldsymbol{r}).$$

In particular, $v^{(3),*}$ approaches a finite limit for all w, I, S, C, t and not $\pm\infty$, as one might expect.

Proof. We start from equation (5.25). The third component of the third term on the right hand side is $O(\varepsilon)$, whereas the third component of the second term is $-\overline{\boldsymbol{a}}^T\overline{\boldsymbol{\Psi}}_0(\boldsymbol{\mu} - \boldsymbol{r}) + O(\varepsilon)$. Noticing that $\frac{C\partial_C g}{g} = \beta \cdot \frac{\partial_x G}{G}$ and $\beta = -\varepsilon\eta^2\varphi_0 + O(\varepsilon^2)$ and furthermore using the result from lemma 15, we can see that the third component of the first term is

$$\eta^2 L(x,t) + O(\varepsilon).$$

This proves the lemma. □

5.3 Valuation of Credit Derivatives

In this section we examine how prices of defaultable bond and CDS rates change under model misspecification. Since the computations are similar to those in sections 3.4 and 3.5, we treat the two kinds of credit derivatives simultaneously.

Under a measure $Q \sim \mathbb{P}$ we assume that the wealth process has the dynamics

$$dW_t = \begin{cases} \Big[((\boldsymbol{\mu} - \boldsymbol{r})^T + (v_t^{(1)}, v_t^{(2)}))\boldsymbol{\pi}_t + r\ W_t + \epsilon AF\Big]\ dt + \\ \qquad\qquad\qquad\qquad + \pi_t^{(1)}\sigma_1\ dB_t^{(1)} + \pi_t^{(2)}\sigma_2\ dB_t^{(2)}, & t < \tau_1, \\ \Big[(\mu_1 - r + v_t^{(1)})\pi_t^{(1)} + r\ W_t + \epsilon AF\Big]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & \tau_1 < t < \tau_2, \\ \Big[(\mu_1 - r + v_t^{(1)})\pi_t^{(1)} + r\ W_t\Big]\ dt + \pi_t^{(1)}\sigma_1 dB_t^{(1)}, & t > \tau_2, \end{cases}$$

subject to

$$\begin{aligned} W_{\tau_1} &= W_{\tau_1^-} + R_2 \cdot \mathbb{I}\{\tau_1 = T\}, \\ W_{\tau_2} &= W_{\tau_2^-} + R_1 \cdot \mathbb{I}\{\tau_2 < T\} + R_2 \cdot \mathbb{I}\{\tau_2 = T\}. \end{aligned}$$

Here R_1 is a random payment independent of the driving Brownian motions, and R_2 is a deterministic payment. The choice $\epsilon = 0$, $R_1 = RF$ (R=recovery), $R_2 = F$ corresponds to an investment in the defaultable bond, whereas $\epsilon = \pm 1$, $R_1 = -\epsilon(1 - R)F$, $R_2 = 0$ corresponds to the investment of the seller/buyer of credit protection.

5.3.1 Valuation in the Distressed Regime

In the distressed regime we assume that the only model uncertainty comes from the drift of the tradable asset I. While it would be both desirable and realistic to incorporate uncertainty in the hazard rate κ as well, we choose not do so here due to analytical tractability. In certain cases this can be somewhat justified by assuming that the default probability of the firm can be estimated fairly well from past data (e.g. from firms within the same market sector and of similar size). Furthermore we also assume complete model certainty in the distribution of the recovery rate.

We define the value function $\overline{V}$ as the solution of the optimization problem

$$\overline{V}(w,I,t) = \sup_{\pi\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2}\int_t^{\tau_d\wedge T} \overline{V}(W_s,I_s,s)\ \phi\ v_s^2\ ds + \right.$$
$$\left. + \frac{1}{2}\int_{\tau_d\wedge T}^{T} V(W_s,s)\ \phi\ v_s^2\ ds \ \middle|\ W_s = w,\ I_s = I \right].$$

This definition is the analog of the definition of V in section 5.2.1. The corresponding HJB equation for $\overline{V}$ is

$$\begin{cases} \partial_t \overline{V} + \epsilon AF\ \partial_w \overline{V} + \sup_{\pi\in\mathbb{R}} \inf_{v\in\mathbb{R}} \left\{ \mathcal{L}^{\pi,v}\ \overline{V} + \frac{1}{2}\ \overline{V}\phi v^2 \right\} + \\ \qquad\qquad +\kappa \left[V(w+\tilde{R}_t,\ t) - \overline{V} \right] = 0\ , \\ \qquad\qquad V(w,T) = u(w+R_2)\ , \quad w\in\mathbb{R}\ , \end{cases} \tag{5.26}$$

where

$$\mathcal{L}^{\pi,v}\overline{V} = (rw + (\mu_1 - r + v)\pi)\ \partial_w \overline{V} + \frac{1}{2}\pi^2\sigma_1^2\ \partial_{ww}\overline{V}$$

and (as before)

$$\tilde{R}_t = \frac{1}{\gamma\epsilon e^{r(T-t)}} \log \mathbb{E}\ e^{\gamma\epsilon R_1 e^{r(T-t)}}.$$

Factoring out wealth by writing $\overline{V}(w,t) = u(we^{r(T-t)})\ \overline{g}(t)$ leads to the following ODE for $\overline{g}$:

$$\begin{cases} \overline{g}' + \inf_{\pi\in\mathbb{R}} \sup_{v\in\mathbb{R}} \left\{ [\epsilon AF + (\mu_1 - r + v)\pi]\ a_t\overline{g} + \frac{1}{2}\sigma_1^2\pi^2 a_t^2\overline{g} + \frac{1}{2}\overline{g}\ \phi v^2 \right\} + \\ \qquad\qquad +\kappa\ \left[e^{-\frac{1}{2}\overline{\lambda}^2(T-t) - \tilde{R}_1 a_t} - \overline{g} \right] = 0, \\ \qquad\qquad \overline{g}(T) = e^{-\gamma R_2}, \end{cases}$$

Carrying out the optimization on the left hand side like in section 5.2.1 leads to

$$\begin{cases} \overline{g}' - \left(\kappa + \frac{1}{2}\overline{\lambda}^2 - \epsilon AF a_t \right)\ \overline{g} + \kappa e^{-\frac{1}{2}\overline{\lambda}^2(T-t) - \tilde{R}_1 a_t} = 0, \\ \qquad\qquad \overline{g}(T) = e^{-\gamma R_2}, \end{cases}$$

which can be solved in the usual way. Things become easy by noticing that for the defaultable bond ($\epsilon = 0$, $R_1 = RF$, $R_2 = F$) the equation above is the same as equation

(2.2), its analog for the completely specified case, only with λ replaced by $\overline{\lambda}$. Similarly, for the CDS we get the same equation as (2.6), where again λ is replaced by $\overline{\lambda}$.

It is interesting to see that since the indifference price of the bond as well as the indifference CDS rates do not depend on $\overline{\lambda}$, they are exactly the same as in the case with complete model certainty. This fact makes it even more interesting to check whether and how uncertainty on κ influences the prices of credit derivatives. This however is material for future research.

Valuation in the Healthy Regime

In analogy to the previous sections we define the value function $\overline{U}(w, P, S, C, t)$ to be the solution of the optimization problem

$$\begin{aligned}\overline{U}(w, I, S, C, t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \Bigg[u(W_T) + \frac{1}{2}\int_t^{\tau_h\wedge T} \overline{U}(W_s, I_s, S_s, C_s, s)\ \boldsymbol{v}_s^T \boldsymbol{\Phi} \boldsymbol{v}_s\ ds\ + \\ +\frac{1}{2}\int_{\tau_h\wedge T}^{\tau_d\wedge T} \overline{V}(W_s, s)\ \phi v_s^2\ ds\ + \frac{1}{2}\int_{\tau_d\wedge T}^{T} V(W_s, s)\ \phi v_s^2\ ds\ \Bigg| \\ \Bigg|\ W_t = w,\ I_t = I,\ S_t = S,\ C_t = C \Bigg]. \quad (5.27)\end{aligned}$$

Here $\boldsymbol{\Phi}$ is the same matrix as in the investment problem in section 5.2.2.

Assuming that $\overline{U}$ is independent of I and S, the corresponding HJB equation is

$$\begin{cases} \partial_t \overline{U} + \epsilon A F\ \partial_w \overline{U} + \sup\limits_{\boldsymbol{\pi}\in\mathbb{R}^2} \inf\limits_{\boldsymbol{v}\in\mathbb{R}^3} \mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}} \overline{U} = 0, \\ \overline{U}(w, C, T) = u(w + R_2) \\ \overline{U}(w, D, t) = \overline{V}(w, t). \end{cases}$$

The operator $\mathcal{L}^{\boldsymbol{\pi},\boldsymbol{v}}$ is the same as in (5.16). As in section 3.5, we make an ansatz of the

form $\overline{U}(w,C,t) = u(we^{r(T-t)})\ e^{\psi(T-t)}\ \overline{h}(C,t)$ with $\psi(\tau) = -\epsilon\gamma\frac{A}{r}e^{r\tau}$ leading to the PDE

$$\begin{cases} \partial_t\overline{h} + \inf\limits_{\boldsymbol{\pi}\in\mathbb{R}^2}\ \sup\limits_{\boldsymbol{v}\in\mathbb{R}^3} \Big\{ \boldsymbol{\pi}^T\left(\boldsymbol{\mu} - \boldsymbol{r} + (v^{(1)}, v^{(2)})^T\right) a_t\overline{h} + \dfrac{1}{2}\boldsymbol{\pi}^T\boldsymbol{\Omega}\boldsymbol{\pi}\ a_t^2\overline{h} + \\ +(\nu + v^{(3)})C\ \partial_C\overline{h} + \boldsymbol{\pi}^T\boldsymbol{\omega}\ a_tC\ \partial_C\overline{h} + \dfrac{1}{2}\eta^2C^2\ \partial_{CC}\overline{h} + \dfrac{1}{2}\overline{h}\ \boldsymbol{v}^T\boldsymbol{\Phi}\boldsymbol{v} \Big\} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overline{h}(C,T) = e^{-\gamma R_2}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overline{h}(D,t) = e^{-\psi(T-\tau)}\cdot\overline{g}(t). \end{cases}$$

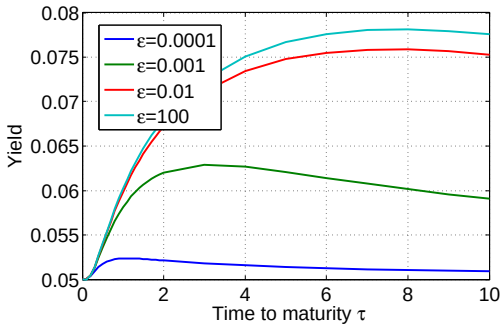
Note that this equation is the same as (5.17), only with different boundary conditions. To solve it, we can therefore make an analogous substitution to get a linear equation. More specifically, we let $\overline{h}(C,t) = \overline{G}^{\beta}(\ln\frac{C}{D},\ T-t)\cdot e^{-\frac{1}{2}\overline{\Lambda}^2(T-t)}$ with $\overline{\Lambda}$, β as in equation (5.21) to get

$$\begin{cases} -\partial_\tau\overline{G} + \widetilde{\nu}\ \partial_x\overline{G} + \dfrac{1}{2}\eta^2\ \partial_{xx}\overline{G} = 0, \\ \qquad\qquad \overline{G}(0,\tau) = e^{-\frac{1}{2\beta}(\overline{\Lambda}^2-\overline{\lambda}^2)\tau - \frac{\psi(\tau)}{\beta}}, \\ \qquad\qquad \overline{G}(x,0) = e^{-\gamma R_2/\beta} \end{cases}$$

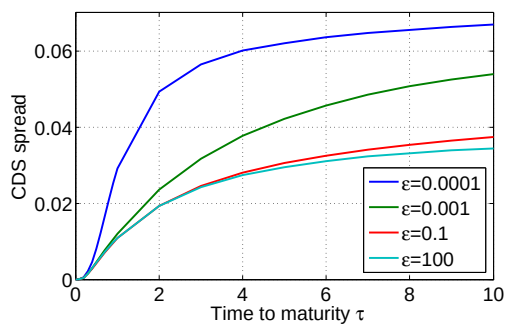
with $\widetilde{\nu}$ as in (5.21). As in the distressed regime, we find the same equation as in the fully specified cases in sections 3.4 and 3.5, only now with the new parameters $\widetilde{\nu}$ and β, and with λ and Λ replaced by $\overline{\lambda}$ and $\overline{\Lambda}$. Once again the MEMM adjusted drift that appeared in the fully specified case disappears and is replaced by a model uncertainty version.

As for the completely specified case, we plot the bond yields as well as the seller's and buyer's CDS spreads. For the following plots we have made specific choices for the negative scalar ϕ_0 and the negative definite matrix $\boldsymbol{\Phi}_0$. This choice is explained in detail at the end of Appendix 5.B. We then let $\phi = \varepsilon\phi_0$, $\boldsymbol{\Phi} = \varepsilon\boldsymbol{\Phi}_0$ for different values of ε. In Figure 5.1 we plot the resulting yields and buyer/seller CDS spreads as the uncertainty varies. The case $\varepsilon = 100$ corresponds to almost complete model certainty. For this case we get almost the same yields and CDS rates as in the completely specified case (for $C_0 = 1.05$). With increasing model uncertainty we observe that the bond yields and seller's CDS rates increase, while the buyer's CDS rates decrease. This is what we intuitively expect. Furthermore, there appears to be more flexibility in the shapes of the

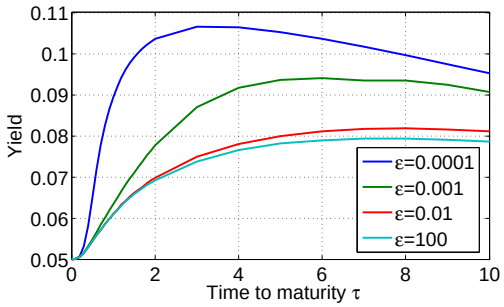
resulting CDS spreads when compared with those in the fully specified case.



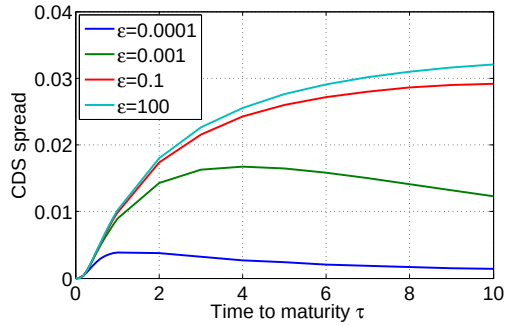
(a) Seller's yields



(b) Seller's CDS spreads



(c) Buyer's Yields



(d) Buyer's CDS spreads

Figure 5.1: The effect of model misspecification on yields and CDS spreads. The parameters for the measure reference measure $\mathbb{P}$ are as for the plots in sections 3.4 and 3.5. The initial CWI was set at $C_0 = 1.05$, risk aversion is $\gamma F = 0.1$. The uncertainty scalar $\phi = \varepsilon\phi_0$ and matrix $\mathbf{\Phi} = \varepsilon\mathbf{\Phi}_0$ with ϕ_0 and $\mathbf{\Phi}_0$ reported in Appendix .

5.A Appendix: Verification Theorem

We assume the same model as in sections 5.2 and 5.3. Let $\tau_1 \triangleq \tau_h \wedge T$ and $\tau_2 \triangleq \tau_d \wedge T$. Furthermore we assume that under a measure $Q \sim \mathbb{P}$, the dynamics of I, S, C are

$$\begin{aligned} dI_t &= I_t \left[\left(\mu_1 + v_t^{(1)} \right) \, dt + \sigma_1 \, dB_t^{(1)} \right], \\ dS_t &= S_t \left[\left(\mu_2 + v_t^{(2)} \right) \, dt + \sigma_2 \, dB_t^{(2)} \right], \\ dC_t &= C_t \left[\left(\nu + v_t^{(3)} \right) \, dt + \eta \, dB_t^{(3)} \right], \end{aligned}$$

and the wealth process W_t has the dynamics

$$dW_t = \begin{cases} \left[\left((\boldsymbol{\mu} - \boldsymbol{r})^T + (v_t^{(1)}, v_t^{(2)}) \right) \boldsymbol{\pi}_t + r \, W_t + \epsilon A F \right] \, dt + \pi_t^{(1)} \sigma_1 \, dB_t^{(1)} + \pi_t^{(2)} \sigma_2 \, dB_t^{(2)}, & t < \tau_1, \\ \left[\left(\mu_1 - r + v_t^{(1)} \right) \pi_t^{(1)} + r \, W_t + \epsilon A F \right] \, dt + \pi_t^{(1)} \sigma_1 \, dB_t^{(1)}, & \tau_1 < t < \tau_2, \\ \left[\left(\mu_1 - r + v_t^{(1)} \right) \pi_t^{(1)} + r \, W_t \right] \, dt + \pi_t^{(1)} \sigma_1 \, dB_t^{(1)}, & t > \tau_2, \end{cases}$$

subject to

$$\begin{aligned} W_{\tau_1} &= W_{\tau_1^-} + R_2 \cdot \mathbb{I}\{\tau_1 = T\}, \\ W_{\tau_2} &= W_{\tau_2^-} + R_1 \cdot \mathbb{I}\{\tau_2 < T\} + R_2 \cdot \mathbb{I}\{\tau_2 = T\}. \end{aligned}$$

As in previous sections, we often write v_t instead of $v_t^{(1)}$. The interpretation of A, R_1, R_2 is the same as section 5.3. The value functions U_1, U_2 and U_3 are defined as solutions of the equations

$$U_1(w, I, t) = \sup_{\pi \in \mathcal{A}} \inf_{Q \sim \mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2} \int_t^T U_1(W_s, I_s, s) \, \phi \, v_s^2 \, ds \;\middle|\; W_t = w, \; I_t = I, \; t > \tau_d \right], \tag{5.28}$$

$$U_2(w, I, t) = \sup_{\pi \in \mathcal{A}} \inf_{Q \sim \mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2} \int_t^{\tau_2} U_2(W_s, I_s, s) \, \phi \, v_s^2 \, ds \; + \frac{1}{2} \int_{\tau_2}^T U_1(w_s, I_s, s) \, \phi \, v_s^2 \, ds \;\middle| \right.$$
$$\left. \;\middle|\; W_s = w, \; I_s = I, \; \tau_h \leq t < \tau_d \right], \tag{5.29}$$

$$U_3(w,I,S,C,t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \Bigg[u(W_T) + \frac{1}{2}\int_t^{\tau_1} U_3(W_s,I_s,S_s,C_s,s)\ \boldsymbol{v}_s^T \boldsymbol{\Phi} \boldsymbol{v}_s\ ds + \\ + \frac{1}{2}\int_{\tau_1}^{\tau_2} U_2(W_s,I_s,s)\ \phi\ v_s^2\ ds\ + \frac{1}{2}\int_{\tau_2}^{T} U_1(w_s,I_s,s)\ \phi\ v_s^2\ ds \ \Bigg| \\ \Bigg|\ W_t = w,\ I_t = I,\ S_t = S,\ C_t = C,\ t < \tau_h \Bigg]. \quad (5.30)$$

Here $\phi < 0$ is a constant and $\boldsymbol{\Phi} \in \mathbb{R}_{3\times 3}$ is a negative definite matrix.

Corresponding to U_1, U_2, U_3 we consider the HJB equations

$$\begin{cases} \partial_t U_1 + \sup_{\pi\in\mathbb{R}} \inf_{v\in\mathbb{R}} \left\{ \mathcal{L}_1^{\pi,v}\ U_1 + \frac{1}{2} U_1\ \phi\ v^2 \right\} = 0\ , \\ \qquad\qquad U_1(w,T) = u(w)\ , \quad w\in\mathbb{R}\ , \end{cases} \quad (5.31)$$

where

$$\mathcal{L}_1^{\pi,v} U_1 = (rw + (\mu_1 - r + v)\pi)\ \partial_w U_1 + \frac{1}{2}\pi^2\sigma_1^2\ \partial_{ww} U_1,$$

$$\begin{cases} \partial_t U_2 + \epsilon A\ \partial_w U_2 + \sup_{\pi\in\mathbb{R}} \inf_{v\in\mathbb{R}} \left\{ \mathcal{L}_2^{\pi,v}\ U_2 + \frac{1}{2} U_2\ \phi\ v^2 \right\} + \\ \qquad\qquad + \kappa \left[U_1(w+\tilde{R}_1,t) - U_2 \right] = 0\ , \\ \qquad\qquad U_2(w,T) = u(w+R_2)\ , \quad w\in\mathbb{R}\ , \end{cases} \quad (5.32)$$

where

$$\mathcal{L}_2^{\pi,v} U_2 = (rw + (\mu_1 - r + v)\pi)\ \partial_w U_2 + \frac{1}{2}\pi^2\sigma_1^2\ \partial_{ww} U_2, \quad \tilde{R}_1 = \frac{1}{\gamma\epsilon e^{r(T-t)}} \log \mathbb{E}\ e^{\gamma\epsilon R_1 e^{r(T-t)}},$$

and

$$\begin{cases} \partial_t U_3 + \epsilon A \partial_w U_3 + \sup_{\boldsymbol{\pi}\in\mathbb{R}^2} \inf_{\boldsymbol{v}\in\mathbb{R}^3} \mathcal{L}_3^{\boldsymbol{\pi},\boldsymbol{v}} U_3 = 0, \\ \qquad U_3(w,C,T) = u(w+R_2), \quad w\in\mathbb{R},\ C > D, \\ \qquad U_3(w,D,t) = U_2(w,t), \quad w\in\mathbb{R},\ t\in[0,T], \end{cases} \quad (5.33)$$

with

$$\mathcal{L}_3^{\boldsymbol{\pi},\boldsymbol{v}} U_3 \ = \ \left[rw + \boldsymbol{\pi}^T \left((\boldsymbol{\mu} - \boldsymbol{r}) + (v^{(1)}, v^{(2)})^T \right) \right]\ \partial_w U_3 + \frac{1}{2}\boldsymbol{\pi}^T \boldsymbol{\Omega}\boldsymbol{\pi}\ \partial_{ww} U_3 + (\nu_1 + v^{(3)}) C\ \partial_C U_3 + \\ + \boldsymbol{\pi}^T \boldsymbol{\omega} C\ \partial_{wC} U_3 + \frac{1}{2}\eta^2 C^2\ \partial_{CC} U_3 + \frac{1}{2}\ U_3\ \boldsymbol{v}^T \boldsymbol{\Phi} \boldsymbol{v}.$$

Let $\overline{\mathbf{\Omega}}$ be the variance-covariance matrix of I, S, C (in contrast to $\mathbf{\Omega}$, the variance-covariance matrix of I, S). In analogy to the completely specified case, the following verification theorem holds:

Theorem 19. *(a) Suppose that there exists a function $H_1 = H_1(w,t)$ that is a solution of (5.31). Furthermore, suppose that for each $(w,t) \in \mathbb{R}\times[0,T]$ there exist $\pi^M = \pi^M(w,t) \in \mathbb{R}$, $v^M = v^M(w,t) \in \mathbb{R}$ such that*

$$\mathcal{L}_1^{\pi^M, v^M} H_1 = \sup_{\pi\in\mathbb{R}} \inf_{v\in\mathbb{R}} \mathcal{L}_1^{\pi,v} H_1. \tag{5.34}$$

Assume that the trading strategy defined by (5.34) is admissible and that there exists a measure $Q^M \sim \mathbb{P}$ under which I has the dynamics

$$dI_t = I_t \ \left[(\mu_1 + v_t^M)\ dt + \sigma_1\ dB_t^{(1)}\right].$$

Then $H_1(w,t)$ is a solution of equation (5.28) for $(w,P,t) \in \mathbb{R}\times[0,\infty)\times[0,T]$.

(b) Suppose there exists a function $H_2 = H_2(w,t)$ that is a solution of (5.32). Furthermore, suppose that for each $(w,t) \in \mathbb{R}\times[0,T]$ there exist $\pi^* = \pi^*(w,t) \in \mathbb{R}$, $v^* = v^*(w,t) \in \mathbb{R}$ such that

$$\mathcal{L}_2^{\pi^*, v^*} H_2 = \sup_{\pi\in\mathbb{R}} \inf_{v\in\mathbb{R}} \mathcal{L}_2^{\pi,v} H_2. \tag{5.35}$$

Assume that the trading strategy $\overline{\pi}_t$ defined by

$$\overline{\pi}_t = \begin{cases} \pi^*(W_t,t), & \tau_h \le t \le \tau_d, \\ \pi_t^M, & t > \tau_d \end{cases}$$

is admissible and that there exists a measure $Q^* \sim \mathbb{P}$ under which I has the dynamics

$$dI_t = I_t \left[(\mu_1 + \overline{v}_t)\ dt + \sigma_1\ dB_t^{(1)}\right],$$

where

$$\overline{v}_t = \begin{cases} v^*, & t < \tau_d, \\ v^M, & t \ge \tau_d. \end{cases}$$

Then H_2 is a solution of (5.29).

(c) Suppose there exists a function $H_3 = H_3(w, C, t)$ which solves (5.33), and suppose that for each $(w, C, t) \in \mathbb{R} \times (D, \infty) \times [0, T]$ there exist $\boldsymbol{\pi}^{**} = \boldsymbol{\pi}^{**}(w, C, t) \in \mathbb{R}^2$ and $\boldsymbol{v}^{**} = \boldsymbol{v}^{**}(w, C, t)$ such that

$$\mathcal{L}_3^{\boldsymbol{\pi}^{**}, \boldsymbol{v}^{**}} H_3 = \sup_{\boldsymbol{\pi} \in \mathbb{R}^2} \inf_{\boldsymbol{v} \in \mathbb{R}^3} \mathcal{L}_3^{\boldsymbol{\pi}, \boldsymbol{v}} H_3. \tag{5.36}$$

Assume that the trading strategy defined by

$$\overline{\boldsymbol{\pi}}_t = \begin{cases} \boldsymbol{\pi}^{**}(W_t, C_t, t), & t < \tau_h, \\ (\overline{\pi}_t, 0), & t \geq \tau_h, \end{cases}$$

is admissible and that there exists a measure $Q^{**} \sim \mathbb{P}$ under which I, S, C have the drift adjustments $\boldsymbol{v}_t^{**}$ up to time $\tau_h \wedge T$, and furthermore I has drift adjustment $\overline{v_t}$ between $\tau_h \wedge T$ and T. Then H_3 is a solution of (5.30).

Proof. We start with part (a). Note that this part is very similar to the verification theorem of the standard Merton investment problem. Since our optimization problem is somewhat different and non-standard, we give the proof anyway.

Consider the measure Q^M and let $\pi \in \mathcal{A}$ be any admissible strategy. Working under Q^M, we get from Ito's lemma

$$H_1(W_T, T) = H_1(w, t) + \int_t^T (\partial_t H_1 + \mathcal{L}_1^{\pi, v^M} H_1)\, ds + \int_t^T \pi_s \sigma_1\, \partial_w H_1\, dB_s^{(1)}.$$

Using the facts that H_1 is a solution of (5.31) and that $\pi \in \mathcal{A}$ is an arbitrary strategy, we always have $\partial_t H_1 + \mathcal{L}^{\pi, v^M} H_1 \leq -\frac{1}{2} H_1 \phi (v^M)^2$. Taking expectations therefore leads to

$$\mathbb{E}_t^{Q^M}\, H_1(W_T, T) \leq H_1(w, t) - \mathbb{E}_t^{Q^M} \left[\frac{1}{2} \int_t^T H_1(W_s, s)\, \phi\, (v_s^M)^2\, ds \right].$$

Using $H_1(W_T, T) = u(W_t)$ we get

$$H_1(w, t) \geq \mathbb{E}_t^{Q^M} \left[u(W_T) + \frac{1}{2} \int_t^T H_1(W_s, s)\, \phi\, (v_s^M)^2\, ds \right].$$

Since this inequality holds for all admissible trading strategies, it follows that

$$\begin{aligned} H_1(w,t) &\geq \sup_{\pi\in\mathcal{A}} \mathbb{E}_t^{Q^M}\left[u(W_T) + \frac{1}{2}\int_t^T H_1(W_s,s)\ \phi\ (v_s^M)^2\ ds\right] \\ &\geq \sup_{\pi\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}_t^{Q}\left[u(W_T) + \frac{1}{2}\int_t^T H_1(W_s,s)\ \phi\ v_s^2\ ds\right]. \end{aligned} \tag{5.37}$$

Now fix the strategy π_t^M and let $Q \sim \mathbb{P}$ be any equivalent measure. Let P_t have drift $\mu_1 + v_t$ under Q. Since we always have $\partial_t H_1 + \mathcal{L}^{\pi^M,v} H_1 \geq -\frac{1}{2} H_1 \phi v_s^2$, a similar argument as above leads to

$$\mathbb{E}_t^Q\ H_1(W_T,T) \geq H_1(w,t) - \mathbb{E}_t^Q\left[\frac{1}{2}\int_t^T H_1(W_s,s)\ \phi\ v_s^2\ ds\right],$$

and hence

$$H_1(w,t) \leq \mathbb{E}_t^Q\left[u(W_T) + \frac{1}{2}\int_t^T H_1(W_s,s)\ \phi\ v_s^2\ ds\right].$$

Since this relation holds for any $Q \sim \mathbb{P}$, we get

$$\begin{aligned} H_1(w,t) &\leq \inf_{Q\sim\mathbb{P}} \mathbb{E}_t^{Q}\left[u(W_T) + \frac{1}{2}\int_t^T H_1(W_s,s)\ \phi\ v_s^2\ ds\right] \\ &\leq \sup_{\pi\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}_t^{Q}\left[u(W_T) + \frac{1}{2}\int_t^T H_1(W_s,s)\ \phi\ v_s^2\ ds\right]. \end{aligned} \tag{5.38}$$

The theorem then follows from inequalities (5.37) and (5.38). For later and in analogy to the verification theorem for the completely specified case we note that for the strategy π^M and the measure Q^M we get equality everywhere, i.e. $H_1(w,t) = \mathbb{E}^{Q^M}\left[u(W_T^{\pi^M} + \frac{1}{2}\int_t^T H_1(W_s^{\pi^M},s)\ \phi\ (v_s^M\right.$

Now we prove part (b). Let π_t be an arbitrary trading strategy and consider the measure Q^*. Then as in part (a), Ito's lemma yields

$$H_2(W_\tau,\tau) = H_2(w,t) + \int_t^\tau (\partial_t H_2 + \epsilon A \partial_w H_2 + \mathcal{L}_2^{\pi,v^*} H_2)\ ds + \int_t^\tau \pi_s \sigma_1\ \partial_w H_2\ dB_s^{(1)},$$

where we have written τ instead of τ_2. Using the fact that H_2 solves equation (5.32), we have $\partial_t H_2 + \epsilon A\ \partial_w H_2 + \mathcal{L}_2^{\pi,v^*} H_2 \leq -\frac{1}{2} U_2 \phi v^2 - \kappa\left[U_1(w + \tilde{R}_1,\ t) - H_2\right]$. Taking expectations we therefore get

$$\begin{aligned} H_2(w,t) \geq& \mathbb{E}_t^{Q^*}\left[H_2(W_{\tau^-},\tau^-)\right] + \\ &+ \mathbb{E}_t^{Q^*}\left[\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ (v_s^*)^2\ ds + \int_t^\tau \kappa\left[U_1(W_s + \tilde{R}_1,\ t) - H_2(W_s,s)\right]\ ds\right]. \end{aligned}$$

As in appendix 2.A, we have

$$\mathbb{E}_t^{Q^*}\left[\int_t^\tau \kappa\left[U_1(W_s+\tilde{R}_1,s)-H_2(W_s,s)\right]\ ds\right]=\mathbb{E}_t^{Q^*}\left[\left(U_1(W_{\tau^-}+\tilde{R}_1,\tau)-H_2(W_{\tau^-},\tau^-)\right)\cdot\mathbb{I}\{\tau_d\le T\}\right].$$

Using $\mathbb{E}_t^{Q^*}\left[U_1(W_{\tau^-}-\epsilon\tilde{R}_1,\tau^-)\cdot\mathbb{I}\{\tau_d\le T\}\right]=\mathbb{E}_t^{Q^*}\left[\ U_1(W_\tau,\tau)\cdot\mathbb{I}\{\tau_d\le T\}\right]$, the inequality above becomes

$$H_2(w,t)\ge\mathbb{E}_t^{Q^*}\left[U(W_T,T)\cdot\mathbb{I}\{\tau_d>T\}+U_1(W_\tau,\tau)\cdot\mathbb{I}\{\tau_d\le T\}+\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ (v_s^*)^2\ ds\right].$$

From the proof of part (a) we know that

$$U_1(W_\tau,\tau)\ge\mathbb{E}_\tau^{Q^*}\left[u(W_T)+\frac{1}{2}\int_\tau^T U_1(W_s,s)\ \phi\ (v_s^*)^2\ ds\right],$$

so we get

$$H_2(w,t)\ge\mathbb{E}_t^{Q^*}\left[u(W_T)+\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ (v_s^*)^2\ ds\ +\frac{1}{2}\int_\tau^T U_1(w_s,s)\ \phi\ (v_s^*)^2\ ds\right],$$

hence, since π is an arbitrary admissible strategy,

$$\begin{aligned}H_2(w,t)&\ge\sup_{\pi\in\mathcal{A}}\mathbb{E}_t^{Q^*}\left[u(W_T)+\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ (v_s^*)^2\ ds\ +\frac{1}{2}\int_\tau^T U_1(W_s,s)\ \phi\ (v_s^*)^2\ ds\right]\\&\ge\sup_{\pi\in\mathcal{A}}\inf_{Q\sim\mathbb{P}}\mathbb{E}_t^{Q}\left[u(W_T)+\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ v_s^2\ ds\ +\frac{1}{2}\int_\tau^T U_1(W_s,s)\ \phi\ v_s^2\ ds\right].\end{aligned}\tag{5.39}$$

Now fix the strategy $\overline{\pi}_t$ and let $Q\sim\mathbb{P}$ be any equivalent measure. Then from arguments analog to those above and in part (a), it follows that

$$\begin{aligned}H_2(w,t)&\le\inf_{Q\sim\mathbb{P}}\mathbb{E}_t^{Q}\left[u(W_T)+\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ v_s^2\ ds\ +\frac{1}{2}\int_\tau^T U_1(W_s,s)\ \phi\ v_s^2\ ds\right]\\&\le\sup_{\pi\in\mathcal{A}}\inf_{Q\sim\mathbb{P}}\mathbb{E}_t^{Q}\left[u(W_T)+\frac{1}{2}\int_t^\tau H_2(W_s,s)\ \phi\ v_s^2\ ds\ +\frac{1}{2}\int_\tau^T U_1(W_s,s)\ \phi\ v_s^2\ ds\right].\end{aligned}\tag{5.40}$$

Them the claim follows from inequalities (5.39) and (5.40). Furthermore, for $\pi=\overline{\pi}$ and $Q=Q^*$ we get equality everywhere, i.e. we have

$$H_2(w,t)=\mathbb{E}_t^{Q^*}\left[u(W_T^{\overline{\pi}})+\frac{1}{2}\int_t^{\tau_2} H_2(W_s^{\overline{\pi}},s)\ \phi\ (v_s^*)^2\ ds+\frac{1}{2}\int_{\tau_2}^T U_1(W_s^{\overline{\pi}},s)\ \phi\ (v_s^*)^2\ ds\ \right].$$

The proof of part (c) is analogous.

□

5.B Appendix: Motivation for the Definition of $\mathbf{\Phi}$

In this section we give the defintion and its motivation of the matrix $\mathbf{\Phi}$ in section 5.2. Suppose that instead of equation (5.5), the definition for U is

$$\begin{aligned} U(w, I, S, C, t) &\triangleq \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \left[u(W_T) + k \log \frac{dQ}{d\mathbb{P}} \right] \\ &= \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \mathbb{E}^Q \left[u(W_T) + \frac{1}{2} k \int_t^T \boldsymbol{v}_s \overline{\mathbf{\Omega}}^{-1} \boldsymbol{v}_s \, ds \right], \quad k > 0, \end{aligned} \tag{5.41}$$

where $\overline{\mathbf{\Omega}}$ is the covariance matrix of I, S and C. This is the analog of the original definition in Anderson, Hansen, and Sargent (2000). Following the idea in Uppal and Wang (2003), we would like to modify this definition to take into account different levels of uncertainty corresponding to the marginal distributions of different subsets of $\{I_t,\ S_t,\ C_t\}$.

Let $A_1, \ldots, A_7$ be the different non-empty subsets of $\{I,\ S,\ C\}$. Corresponding to a subset A_i, we start with the matrix $\widetilde{\mathbf{\Omega}}_i^{-1} \in \mathbb{R}^{|A_i|\times|A_i|}$, the inverse of the variance-covariance matrix of the elements contained in A_i. Then we define the matrix $\mathbf{\Omega}_i^{-1} = \begin{pmatrix} \omega_{II}^{(i)} & \omega_{IS}^{(i)} & \omega_{IC}^{(i)} \\ \omega_{SI}^{(i)} & \omega_{SS}^{(i)} & \omega_{SC}^{(i)} \\ \omega_{CI}^{(i)} & \omega_{CS}^{(i)} & \omega_{CC}^{(i)} \end{pmatrix}$ as follows (following Uppal and Wang (2003)):

- Let $X, Y \in \{I,\ S,\ C\}$. If $X \notin A_i$ or $Y \notin A_i$, then $\omega_{XY}^{(i)} = 0$.
- If we delete all rows and columns of $\mathbf{\Omega}_i^{-1}$ indexed by elements not contained in A_i, the resulting matrix is $\widetilde{\mathbf{\Omega}}_i^{-1}$. [1]

Then $\mathbf{\Phi}$ is defined as linear combination

$$\mathbf{\Phi} \triangleq \sum_{i=1}^{7} \alpha_i \mathbf{\Omega}_i^{-1}, \tag{5.42}$$

[1] Example: Let $A_i = \{I_t,\ C_t\}$. If $\widetilde{\mathbf{\Omega}}_i = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\eta \\ \rho\sigma_1\eta & \eta^2 \end{pmatrix}$ is the variance-covariance matrix of I_t and C_t, then $\widetilde{\mathbf{\Omega}}_i^{-1} = \frac{1}{\sigma_1^2\eta^2(1-\rho^2)} \begin{pmatrix} \eta^2 & -\rho\sigma_1\eta \\ -\rho\sigma_1\eta & \sigma_1^2 \end{pmatrix}$ and hence $\mathbf{\Omega}_i^{-1} = \frac{1}{\sigma_1^2\eta^2(1-\rho^2)} \begin{pmatrix} \eta^2 & 0 & -\rho\sigma_1\eta \\ 0 & 0 & 0 \\ -\rho\sigma_1\eta & 0 & \sigma_1^2 \end{pmatrix}$.

where the α_i are *non-positive* weights. A large $|\alpha_i|$ corresponds to a high level of certainty for the marginal distribution of the corresponding subset, whereas a small $|\alpha_i|$ corresponds to a high level of uncertainty. Note that as a negative linear combination of positive semi-definite matrices, $\mathbf{\Phi}$ is negative semi-definite. If the weight corresponding to the uncertainty in the joint distribution of I_t, S_t, C_t is negative, then $\mathbf{\Phi}$ is strictly negative definite.

Given a measure $Q \sim \mathbb{P}$ and a non-empty subset $A_i \subseteq \{I_t,\ S_t,\ C_t\}$, let $\mathbb{P}_i$, Q_i be the induced measures for the marginal distribution of the elements in A_i. Then an appropriate modification of (5.41) is

$$U(w, I, S, C, t) \triangleq \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \left\{ \mathbb{E}^Q\left[u(W_T)\right] + \sum_{i=1}^{7} \alpha_i \mathbb{E}^{Q_i}\left[\log \frac{dQ_i}{d\mathbb{P}_i}\right] \right\}, \quad w_i < 0. \tag{5.43}$$

For a given subset A_i, let $\boldsymbol{v}_t^{A_i} \in \mathbb{R}^{|A_i|}$ be the vector obtained from $\boldsymbol{v}_t \in \mathbb{R}^3$ by deleting the components that correspond to elements not contained in A_i. Then it is easy to see that

$$\mathbb{E}^{Q_i}\left[\log \frac{dQ_i}{d\mathbb{P}_i}\right] = \mathbb{E}^{Q_i}\left[\frac{1}{2}\int_t^T (\boldsymbol{v}_s^{A_i})^T \widetilde{\mathbf{\Omega}}_i^{-1} \boldsymbol{v}_s^{A_i}\ ds\right] = \mathbb{E}^{Q}\left[\frac{1}{2}\int_t^T \boldsymbol{v}_s^T \mathbf{\Omega}_i^{-1} \boldsymbol{v}_s\ ds\right],$$

so that (5.43) becomes

$$U(w, I, S, C, t) = \sup_{\boldsymbol{\pi}\in\mathcal{A}} \inf_{Q\sim\mathbb{P}} \left[u(W_T) + \frac{1}{2}\int_t^T \boldsymbol{v}_s^T \mathbf{\Phi} \boldsymbol{v}_s\ ds\right].$$

Applying the same modification to this equation as in Maenhout (2004) then leads to the definition in (5.5).

To conclude this section we explain the choices of the matrix $\mathbf{\Phi}_0$ and the scalar ϕ_0 for the plots in section 5.3. For $\mathbf{\Phi}_0$ we chose all the weights α_i to equal -1, i.e.

$$\mathbf{\Phi}_0 = -\sum_{i=1}^{7} \mathbf{\Omega}_i^{-1}.$$

Furthermore, it is reasonable to assume that the level of uncertainty in the marginal distribution of P does not change upon switching from the healthy to the distressed

regime. Consequently, if we assume that

$$\mathbf{\Omega}_1^{-1} = \begin{pmatrix} \sigma_1^{-2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then a suitable choice for the scalar ϕ is

$$\phi = \alpha_1 \sigma_1^{-2}.$$

Therefore we have chosen $\phi_0 = -\sigma_1^{-2}$ for the plots in section 5.3.

Chapter 6

Pricing of a CDS in a Discrete Time Setting

6.1 Introduction and Model

This chapter was the original starting point of this thesis. It is related to the previous chapters in the sense that we use utility indifference pricing to determine CDS rates in a structural framework. It is independent from them in the sense that the framework used is not continuous, but discrete. In contrast to previous sections, the CDS payments are made at discrete points in time. Moreover, default is also monitored at discrete times only.

Suppose we have a market with a risky default-free asset I, a defaultable asset S (the reference entity's stock) which is correlated to I, and a money market account with constant interest rate r, and suppose that we would like to price a credit default swap written on S. We adopt the setup from section 2.1.2, i.e. due to the default risk, the investor chooses not to invest in S.

Let $D > 0$ be a given time-invariant observable threshold, and let a discrete set of times $0 = t_0$, t_1, t_2, ..., $t_n = T$ be given. The firm defaults, if at any of these times, S is

below D. This setup corresponds to the classical Merton model for pricing a defaultable bond. introduced in Merton (1974), in this model the firm's asset value is a geometric Brownian motion, and the company defaults if at maturity T the asset value is below a certain default level, which can be interpreted as the companies debt.

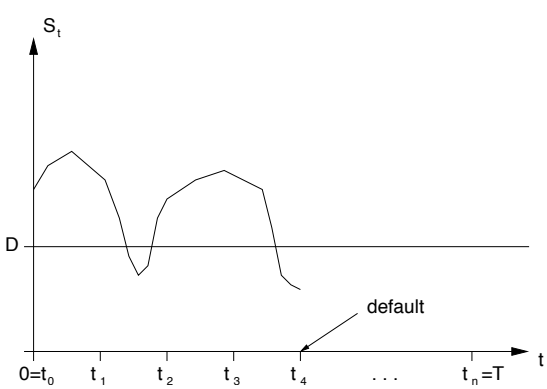
Consequently, we assume that the buyer of a CDS written on the firm makes/receives payments subject to the following rules:

- At time $t_0 = 0$ the buyer makes a payment of A dollars, if $S_0 \geq D$. If $S_0 < D$, no payments are made or received at all.

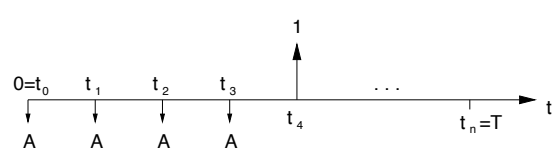
  At time t_i ($i \in \{1, \ldots, n-1\}$) the buyer makes a payment of A dollars, provided that at $t = t_0, \ldots, t_i$ the stock price S_t has been at or above some threshold level D. At time $t_n = T$ the buyer does not make another payment.

- If $S_{t_j} \geq D$ for $j = 0, \ldots, i-1$, but $S_{t_i} < D$, then the buyer receives a one-time payment of 1 dollar at time t_i and does not make any further payments in the future.

The problem is to find an appropriate value of A.



(a) Default model



(b) Payment stream from the buyer's point of view

Note that in this chapter we only consider the buyer's CDS rate. However, it is easy to modify the analysis for the seller as we did in previous chapters. Moreover we assume

that the face value of the firm's defaultable bond is 1 and that the recovery rate is 0. This is mostly for notational purposes and can also easily extended to non-zero or even random recovery rates.

We model I and S as stochastic processes on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. The dynamics of I and S are given by

$$\begin{aligned} dI_t &= I_t \left(\mu \, dt + \sigma \, dB_t^{(1)} \right), \\ dS_t &= S_t \left(\nu \, dt + \eta \, dB_t^{(2)} \right), \end{aligned}$$

where $B_t^{(1)}$ and $B^{(2)}$ are correlated Brownian motions on Ω with correlation ρ $(-1 \leq \rho \leq 1)$. The filtration $(\mathcal{F}_t)$ is the standard filtration generated by $B_t^{(1)}$, $B_t^{(2)}$ and the null sets of Ω.

An trading strategy is given by π_t, the amount invested in I at time t. We demand that π_t satisfy the conditions in definition 2 for $n = 1$ from chapter 1.1 and denote the set of admissible trading strategies by $\mathcal{A}$.

6.2 The Two Investment Problems

We start by maximizing expected utility of terminal wealth when the investor invests in I and the money market only. This is the standard Merton investment problem with one risky asset, and hence by the result from section 1.1 the corresponding value function is

$$V(w,t) = u(w \; e^{r(T-t)}) \; e^{-\frac{1}{2}\lambda^2 (T-t)}, \quad \lambda = \frac{\mu - r}{\sigma}.$$

We now examine the case in which the buyer is invested in one unit of the CDS and in I and the money market with his remaining money. For $\pi \in \mathcal{A}$ we model the wealth

process $\widetilde{W}_t^\pi$ as the unique strong solution of

$$\begin{aligned}
d\widetilde{W}_t &= (r\widetilde{W}_t + (\mu - r)\pi_t)\, dt + \pi_t \sigma\, dB, \quad t \in (t_i, t_{i+1}), \\
\widetilde{W}_{t_i} &= \widetilde{W}_{t_i^-} + \mathbb{I}\left\{S_{t_0} \geq D, \ldots, S_{t_{i-1}} \geq D\right\} \cdot \left(\mathbb{I}\left\{S_{t_i} < D\right\} - A \cdot \mathbb{I}\left\{S_{t_i} \geq D\right\}\right), \\
&\quad i \in \{1, \ldots, n-1\}, \\
\widetilde{W}_T &= \widetilde{W}_{T^-} + \mathbb{I}\left\{S_{t_0} \geq D,\ \ldots,\ S_{t_{n-1}} \geq D,\ S_T < D\right\}.
\end{aligned}$$

By using this definition, we assume that at time $t_0 = 0$, the initial payment of A dollars has already been made.

We now define the value function U by

$$U(w, I, S, t) = \sup_\pi E(u(\widetilde{W}_T) \mid \widetilde{W}_t = w, I_t = I, S_t = S,\ S_{t_j} \geq D \text{ for all } t_j \leq t),$$

i.e. we assume that S has not defaulted yet at time t. Note that this makes U undefined for all (w, I, S, t) such that $t = t_i$ for some i and $S < D$. As in previous chapters, it is clear that U is independent of I, so that we will write $U(w, S, t)$ from now on.

Let $S \geq D$. Then in analogy to previous chapters the indifference price for the periodic payment $A = A(w, I, S)$ is defined by the equation

$$U(w - A, S, 0) = V(w, 0).$$

Assuming that in each open interval (t_i, t_{i+1}), U is continuously differentiable once in t and twice in w, S, we expect U to satisfy the following HJB equation in (t_i, t_{i+1}):

$$\begin{aligned}
U_t + \nu S\ \partial_S U + rw\ \partial_w U + \frac{1}{2}\eta^2 S^2\ \partial_{SS} U + \\
+ \sup_{\pi \in \mathbb{R}} \left\{ \frac{1}{2}\pi^2\sigma^2\ \partial_{ww} U + \pi\left[(\mu - r)\ \partial_w U + \sigma\eta\rho S\ \partial_{Sw} U\right] \right\} &= 0. \qquad (6.1)
\end{aligned}$$

At $t = t_0,\ t_1,\ \ldots,\ T$ however we cannot expect U to be continuous, due to the jumps in $\widetilde{W}$. Therefore one problem is to determine the proper boundary conditions for U at $t = t_0,\ t_1,\ \ldots,\ T$.

Intuitively, we can get the left-hand limit of $U(w, S, t)$ (as $t \to t_j$) in the following way: Suppose that $S_t < D$ and t is close to t_j. Then we expect S to default at t_j, in

which case we will have $\widetilde{W}_{t_j} = \widetilde{W}_{t_j^-} + 1$. Then starting at t_j, we are in the situation of the Merton problem, so that we expect $U(w, S, t) \to V(w + 1, t_j)$ as $t \to t_j^-$.

Similarly, suppose that $S_t > D$ and t is close to t_j. Then we expect S not to default at $t = t_j$, in which case we get $\tilde{W}_{t_j} = \tilde{W}_{t_j^-} - A$, and hence $U(w, P, S, t) \to U(w - A, S, t_j)$ as $t \to t_j^-$.

Similar arguments work to determine $\lim_{t \to T^-} U(w, P, S, t)$, only that in this case the buyer does not make a payment, if $S_T > D$.

The boundary condition for U at $t = t_i$ ($i \in \{1, \ldots, n-1\}$) is therefore given by

$$U(w, S, t_i^-) = \begin{cases} U(w - A, S, t_i), & S > D \\ V(w + 1, t_i), & S < D, \end{cases} \tag{6.2}$$

while at $t = T$ the boundary condition is given by

$$U(w, S, T^-) = \begin{cases} -\frac{1}{\gamma} e^{-\gamma w}, & S > D \\ -\frac{1}{\gamma} e^{-\gamma(w+1)}, & S < D. \end{cases} \tag{6.3}$$

Note that the equations above do not make a statement about the limit $U(w, S, t_i^-)$ for $S = D$. This however does not matter, because $\{S_{t_i} = D\}$ is a set of measure 0.

6.3 Computation of U

The method of finding U is very similar to the previous chapters. We make an ansatz of the form $U(w, S, t) = u(w\ e^{r(T-t)})\ e^{-\frac{1}{2}\lambda^2(T-t)}\ G(s, \tau)$ with new variables

$$s = \log S + \left(\delta - \frac{\eta^2}{2}\right)(T - t), \quad \tau = T - t,$$

with $\delta = \nu - (\mu - r) \cdot \rho \cdot \frac{\eta}{\sigma}$, and additionally define

$$\tau_i = T - t_{n-i} \text{ for } i = 0, \ldots, n.$$

Then $G(s, \tau)$ solves the heat equation

$$\partial_\tau G - \frac{1}{2}\eta^2\ \partial_{ss} G = 0 \tag{6.4}$$

subject to the boundary conditions

$$G(s,0^+) = \begin{cases} 1, & s > \log D, \\ e^{-\gamma(1-\rho^2)}, & s < \log D, \end{cases}$$

and

$$G(s,\tau_i^+) = \begin{cases} e^{A\gamma(1-\rho^2)e^{r\tau_i}} \cdot G(s,\tau_i), & s > \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_i, \\ e^{-\gamma(1-\rho^2)e^{r\tau_i}}, & s < \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_i. \end{cases}$$

Note that due to the discrete default times (instead of a continuous default barrier) we could remove the drift term in equation (6.4).

Using the Feynman-Kac formula the solution of the equation above can be recursively represented as

$$G(s,\tau) = E\ G(s + \eta(Z_\tau - Z_{\tau_i}),\ \tau_i^+), \quad \tau \in (\tau_i, \tau_{i+1}],$$

where Z_t is a Wiener process on Ω. An explicit representation for G will be derived in the next sections.

6.4 Behaviour of the Indifference Price as $\gamma \to 0$

Let $\widetilde{\mathbb{P}}$ be any measure under which S_t has the dynamics

$$S_t = S_0 \cdot e^{\left(\delta - \frac{\eta^2}{2}\right)t + \eta W_t},$$

where W_t is a standard Brownian motion. In particular, $\widetilde{\mathbb{P}}$ can be the minimal martingale measure. In this section we will compute the risk-neutral CDS spread $\tilde{A}$ under $\widetilde{\mathbb{P}}$. We will then show that the CDS spread obtained by the method of indifference pricing converges to $\tilde{A}$ as $\gamma \to 0$.

By comparing the expected discounted payments made and received, it is easy to see

that the risk-neutral CDS spread is given by the equation

$$\begin{aligned} A + \widetilde{\mathbb{P}}(S_{t_1} < D) \cdot (-1)e^{-rt_1} + \widetilde{\mathbb{P}}(S_{t_1} \geq D,\ S_{t_2} < D) \cdot \left[Ae^{-rt_1} - 1 \cdot e^{-rt_2}\right] \\ + \widetilde{\mathbb{P}}(S_{t_1} \geq D,\ S_{t_2} \geq D,\ S_{t_3} < D) \cdot \left[Ae^{-rt_1} + Ae^{-rt_2} - 1 \cdot e^{-rt_3}\right] + \ldots + \\ + \widetilde{\mathbb{P}}(S_{t_1} \geq D, \ldots,\ S_{t_{n-1}} \geq D,\ S_T < D) \cdot \left[Ae^{-rt_1} + \ldots + Ae^{-rt_{n-1}} - 1 \cdot e^{-rt_T}\right] \\ + \widetilde{\mathbb{P}}(S_{t_1} \geq D, \ldots,\ S_T \geq D) \cdot \left[Ae^{-rt_1} + \ldots + Ae^{-rt_{n-1}}\right] = 0. \end{aligned} \tag{6.5}$$

For simplicity of notation, let $a_1, \ldots, a_{n+1}$ denote the probabilities on the left hand side of (6.5), i.e.

$$\begin{aligned} a_1 &= \widetilde{\mathbb{P}}(S_{t_1} < D), \\ a_k &= \widetilde{\mathbb{P}}(S_{t_1} \geq D, \ldots,\ S_{t_{k-1}} \geq D,\ S_{t_k} < D), \quad k = 2, \ldots n, \\ a_{n+1} &= \widetilde{\mathbb{P}}(S_{t_1} \geq D, \ldots,\ S_T \geq D), \end{aligned}$$

and let $x_1, \ldots, x_{n+1}$ denote the corresponding discounted payments, i.e.

$$\begin{aligned} x_1 &= -1 \cdot e^{-rt_1}, \\ x_k &= A \cdot e^{-rt_1} + \ldots + A \cdot e^{-rt_{k-1}} - 1 \cdot e^{-rt_k},\ \ k = 2, \ldots, n, \\ x_{n+1} &= A \cdot e^{-rt_1} + \ldots + A \cdot e^{-rt_{n-1}}. \end{aligned}$$

Then (6.5) becomes

$$A + \sum_{i=1}^{n+1} a_i x_i = 0,$$

where the a_i satisfy $\sum_{i=1}^{n+1} a_i = 1$ and $a_i \geq 0$. For the risk-neutral CDS spread this yields

$$\tilde{A} = \frac{\sum_{j=1}^{n} a_j e^{-rt_j}}{1 + \sum_{j=1}^{n-1} e^{-rt_j} \left(\sum_{i=j+1}^{n+1} a_i\right)}. \tag{6.6}$$

The CDS spread $A_\gamma = A_\gamma(S)$ obtained from indifference pricing is given by the equation

$$A + e^{-rT} \cdot \frac{1}{\gamma(1-\rho^2)} \log G_{A,\gamma}(s,T) = 0, \tag{6.7}$$

with $s = \log S + \left(\delta - \frac{\eta^2}{2}\right) T$. Here $G_{A,\gamma}$ is the function G from the previous sections. However in this section, we want to emphasize its dependence on A and γ.

Theorem 20. *For all $A \geq 0$, $\gamma > 0$, the function $G_{A,\gamma}(s,T)$ has the representation*

$$G_{A,\gamma}(s,T) = \sum_{i=1}^{n+1} a_i e^{\gamma(1-\rho^2)e^{rT}x_i}.$$

The proof is somewhat lengthy and will be presented in appendix 6.A.

For fixed s, let $\tilde{L}(A)$ denote the left hand side of (6.5), and let $L(A,\gamma)$ denote the left hand side of (6.7). Obviously, $\tilde{L}(A)$ is strictly increasing in A.

Lemma 21. *(a) For fixed s, $L(A,\gamma)$ is strictly increasing in both A and γ.*

(b) For fixed s and A,

$$L(A,\gamma) \to \tilde{L}(A) \ \ as \ \ \gamma \to 0.$$

Proof. (a) It is clear that $L(A,\gamma)$ is strictly increasing in A.

To show that $L(A,\gamma)$ is also strictly increasing in γ, we substitute $\beta = \gamma(1-\rho^2)e^{rT}$ and define $f(\beta) = \log \sum_{i=1}^{n+1} a_i e^{\beta x_i}$. Note that $f(\beta) = 0$. We now have to show that $\frac{f(\beta)}{\beta}$ is increasing in β.

f is convex in β, because

$$f'(\beta) = \frac{\sum_{i=1}^{n+1} a_i x_i e^{\beta x_i}}{\sum_{i=1}^{n+1} a_i e^{\beta x_i}}$$

and

$$\begin{aligned} f''(\beta) &= \frac{\left(\sum a_i x_i^2 e^{\beta x_i}\right)\left(\sum a_i e^{\beta x_i}\right) - \left(\sum a_i x_i e^{\beta x_i}\right)^2}{\left(\sum a_i e^{\beta x_i}\right)^2} \\ &= \frac{\sum_{i\neq j} a_i a_j x_i^2 e^{\beta(x_i+x_j)} - \sum_{i\neq j} a_i a_j x_i x_j e^{\beta(x_i+x_j)}}{\left(\sum a_i e^{\beta x_i}\right)^2} \\ &= \frac{\sum_{i<j} e^{\beta(x_i+x_j)} a_i a_j (x_i - x_j)^2}{\left(\sum a_i e^{\beta x_i}\right)^2} \\ &\geq 0. \end{aligned}$$

Since f is convex, the difference quotient $\frac{f(\beta)-f(0)}{\beta-0}$ is increasing in β. This proves the claim.

For (b), we have to show that

$$\lim_{\beta\to 0} \frac{1}{\beta} \log \sum_{i=1}^{n+1} a_i e^{\beta x_i} = \sum_{i=1}^{n+1} a_i x_i.$$

This follows easily from l'Hopital's rule, because

$$\begin{aligned}
\lim_{\beta\to 0} \frac{1}{\beta} \log \sum_{i=1}^{n+1} a_i e^{\beta x_i} &= \lim_{\beta\to 0} \frac{1}{\sum_{i=1}^{n+1} a_i e^{\beta x_i}} \cdot \sum_{i=1}^{n+1} a_i x_i e^{\beta x_i} \\
&= \frac{\sum_{i=1}^{n+1} a_i x_i}{\sum_{i=1}^{n+1} a_i} \\
&= \sum_{i=1}^{n+1} a_i x_i.
\end{aligned}$$

□

Recall that $\tilde{A}$ is the risk-neutral CDS spread. Let A_γ be the CDS spread found by indifference pricing corresponding to γ.

Theorem 22.

$$\lim_{\gamma\to 0} A_\gamma = \tilde{A}.$$

Proof. By definition, A_γ solves $L(A_\gamma, \gamma) = 0$, and $\tilde{A}$ solves $L(\tilde{A}) = 0$. From claim 21 we know that $L(A, \gamma) \downarrow \tilde{L}(A)$ for fixed A as $\gamma \to 0^+$. It follows that A_γ increases as $\gamma \to 0^+$ and that $A_\gamma \leq \tilde{A}$ for all $\gamma > 0$. Hence $\lim_{\gamma\to 0} A_\gamma$ exists and

$$\lim_{\gamma\to 0} A_\gamma \leq \tilde{A}.$$

Let $\hat{A} := \lim_{\gamma\to 0} A_\gamma$ and assume that $\hat{A} < \tilde{A}$. Then $0 = \tilde{L}(\tilde{A}) > L(\hat{A})$. Furthermore, for any $\gamma > 0$ we have $L(\hat{A}, \gamma) > \tilde{L}(A_\gamma, \gamma) = 0$. So we have

$$L(\hat{A}, \gamma) > 0 > \tilde{L}(\hat{A})$$

for all $\gamma > 0$. On the other hand, $\lim_{\gamma\to 0} L(\hat{A}, \gamma) = \tilde{L}(\hat{A})$, i.e. we get a contradiction. This proves that $\lim_{\gamma\to 0} A_\gamma = \tilde{A}$.

□

Now we compute the first-order correction term of A_γ with respect to γ, i.e. we want to write

$$A_\gamma = \tilde{A} + A^{(1)}\gamma + o(\gamma) \qquad (\gamma \to 0).$$

Assuming differentiability of A_γ with respect to γ (which will follow), we have

$$A^{(1)} = \left.\frac{dA}{d\gamma}\right|_{\gamma=0}.$$

For $(A,\gamma) \in \mathbb{R}^2$ consider the function

$$L(A,\gamma) = \begin{cases} A + \frac{1}{\gamma(1-\rho^2)e^{rT}} \log \sum_{i=1}^{n+1} a_i e^{\gamma(1-\rho^2)e^{rT}x_i}, & \gamma \neq 0, \\ A + \sum_{i=1}^{n+1} a_i x_i, & \gamma = 0. \end{cases}$$

If L is continuously differentiable as a function in (A,γ) at $(\tilde{A},0)$ and if $\frac{\partial L}{\partial A}(\tilde{A},0) \neq 0$, then from the implicit function theorem it follows that A_γ is continuously differentiable at $\gamma = 0$ and that

$$\left.\frac{dA}{d\gamma}\right|_{\gamma=0} = -\frac{\frac{\partial L}{\partial \gamma}(\tilde{A},0)}{\frac{\partial L}{\partial A}(\tilde{A},0)}.$$

Hence we now have to show that $\frac{\partial L}{\partial A}$ and $\frac{\partial L}{\partial \gamma}$ both exist in a neighbourhood of $(\tilde{A},0)$ and that they are continuous there.

This is fairly easy to see for $\frac{\partial L}{\partial A}$. We get

$$\begin{aligned} \frac{\partial L}{\partial A}(\tilde{A},0) &= 1 + \sum_{i=1}^{n+1} a_i \left.\frac{\partial x_i}{\partial A}\right|_{A=\tilde{A}} \\ &= 1 + \sum_{j=1}^{n-1} e^{-rt_j} \left(\sum_{i=j+1}^{n+1} a_i \right). \end{aligned}$$

For $\frac{\partial L}{\partial \gamma}(A,0)$ we note that we can write

$$\sum_{i=1}^{n+1} a_i e^{\gamma(1-\rho^2)e^{rT}x_i} = \sum_{k=0}^{\infty} b_k \gamma^k$$

with $b_0 = 1$, $b_1 = (1-\rho^2)e^{rT}\sum_{i=1}^{n+1} a_i x_i$ and $b_2 = \frac{1}{2}(1-\rho^2)^2\left(e^{rT}\right)^2 \sum_{i=1}^{n+1} a_i x_i^2$. Hence for $|\gamma| < 1$ we have

$$\log \sum_{i=1}^{n+1} a_i e^{\gamma(1-\rho^2)e^{rT}x_i} = \sum_{k=0}^{\infty} c_k \gamma^k,$$

where $c_0 = 0$, $c_1 = b_1 = (1-\rho^2)e^{rT}\sum_{i=1}^{n+1} a_i x_i$ and

$$\begin{aligned}
c_2 &= b_2 - \frac{1}{2}b_1^2 \\
&= \frac{1}{2}(1-\rho^2)^2\left(e^{rT}\right)^2\sum_{i=1}^{n+1} a_i x_i^2 - \frac{1}{2}(1-\rho^2)^2\left(e^{rT}\right)\left(\sum_{i=1}^{n+1} a_i x_i\right)^2 \\
&= \frac{1}{2}(1-\rho^2)^2\left(e^{rT}\right)^2\cdot\left[\left(\sum_{i=1}^{n+1} a_i x_i^2\right)\left(\sum_{i=1}^{n+1} a_i\right) - \left(\sum_{i=1}^{n+1} a_i x_i\right)^2\right] \\
&= \frac{1}{2}(1-\rho^2)^2\left(e^{rT}\right)^2\cdot\sum_{i<j} a_i a_j (x_i - x_j)^2.
\end{aligned}$$

Therefore

$$\frac{\partial L}{\partial \gamma}(A,0) = \frac{1}{(1-\rho^2)e^{rT}}c_2 = \frac{1}{2}(1-\rho^2)e^{rT}\cdot\sum_{i<j} a_i a_j (x_i - x_j)^2$$

and

$$A^{(1)} = -\frac{\frac{1}{2}(1-\rho^2)e^{rT}\cdot\sum_{i<j} a_i a_j (\tilde{x_i} - \tilde{x_j})^2}{1 + \sum_{j=1}^{n-1} e^{-rt_j}\left(\sum_{i=j+1}^{n+1} a_i\right)},$$

where $\tilde{x_j} = x_j|_{A=\tilde{A}}$.

Using (6.8) we can rewrite this result as

$$A^{(1)} = -\frac{1}{2\tilde{A}}(1-\rho^2)e^{rT}\cdot\left(\sum_{i<j} a_i a_j (\tilde{x_i} - \tilde{x_j})^2\right)\cdot\left(\sum_{j=1}^{n} a_j e^{-rt_j}\right).$$

6.A Appendix: Proof of Theorem 20

In the previous section we assumed that S_t has the dynamics

$$S_t = S\cdot e^{\left(\delta - \frac{\eta^2}{2}\right)t + \eta W_t}$$

under $\widetilde{\mathbb{P}}$. In terms of the driving Brownian motion, the condition $S_t \geq D$ is equivalent to

$$W_t \geq \frac{\ln\frac{D}{S} - \left(\delta - \frac{\eta^2}{2}\right)t}{\eta}.$$

We can therefore rewrite the $a_1, \ldots, a_{n+1}$ as

$$\begin{aligned} a_1 &= \widetilde{\mathbb{P}}(W_{t_1} < d_1), \\ a_k &= \widetilde{\mathbb{P}}(W_{t_1} \geq d_1,\ \ldots,\ W_{t_{k-1}} \geq d_{k-1},\ W_{t_k} < d_k), \quad k = 2, \ldots, n, \\ a_{n+1} &= \widetilde{\mathbb{P}}(W_{t_1} \geq d_1,\ \ldots,\ W_{t_n} \geq d_n), \end{aligned}$$

where

$$d_k = \frac{\ln \frac{D}{S} - \left(\delta - \frac{\eta^2}{2}\right) t_k}{\eta}, \quad k = 1, \ldots, n.$$

As in previous sections, let $s = \log S + \left(\delta - \frac{\eta^2}{2}\right)(T - t)$, $\tau = T - t$, and define $\tau_i = T - t_{n-i}$ for $i = 0, \ldots, n$. Recall that $G(s, \tau)$ satisfies the heat equation $\partial_\tau - \frac{1}{2}\eta^2 \partial_{ss} G = 0$ subject to the boundary conditions

$$G(s, 0^+) = \begin{cases} 1, & s > \log D, \\ e^{-\gamma(1-\rho^2)}, & s < \log D, \end{cases}$$

and

$$G(s, \tau_i^+) = \begin{cases} e^{A\gamma(1-\rho^2)e^{r\tau_i}} \cdot G(s, \tau_i), & s > \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_i, \\ e^{-\gamma(1-\rho^2)e^{r\tau_i}}, & s < \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_i. \end{cases}$$

If Z_t is a standard Brownian motion on Ω, we can compute $G(s, T)$ recursively by using the relation

$$G(s, \tau_{i+1}) = E\ G(s + \eta(Z_{\tau_{i+1}} - Z_{\tau_i}),\ \tau_i^+).$$

The following lemma gives an explicit formula for $G(s, \tau_m)$ for $m = 1, \ldots, n$ and will prove theorem 20.

Lemma 23. *For $m = 2, \ldots, n$ we have*

$$G(s, \tau_m) = \sum_{i=1}^{m+1} a_i^{(m)}(s) e^{\gamma(1-\rho^2)x_i^{(m)}},$$

where

$$\begin{aligned} x_1^{(m)} &= -1 \cdot e^{r\tau_{m-1}} \\ x_k^{(m)} &= A \cdot e^{r\tau_{m-1}} + \ldots + A \cdot e^{r\tau_{m-k+1}} - 1 \cdot e^{r\tau_{m-k}}, \quad k = 2, \ldots, m, \\ x_{m+1}^{(m)} &= A \cdot e^{r\tau_{m-1}} + \ldots + A \cdot e^{r\tau_1} \end{aligned}$$

and

$$\begin{aligned}
a_1^{(m)}(s) &= \mathbf{P}(Z_{\tau_m} - Z_{\tau_{m-1}} < d_1^{(m)}(s)), \\
a_k^{(m)}(s) &= \mathbf{P}(Z_{\tau_m} - Z_{\tau_{m-1}} \geq d_m^{(1)}(s),\ \ldots,\ Z_{\tau_m} - Z_{\tau_{m-k+1}} \geq d_m^{(k-1)}(s), \\
&\qquad Z_{\tau_m} - Z_{\tau_{m-k}} < d_m^{(k)}(s)), \quad k = 2, \ldots, m, \\
a_{m+1}^{(m)}(s) &= \mathbf{P}(Z_{\tau_m} - Z_{\tau_{m-1}} \geq d_m^{(1)}(s),\ \ldots,\ Z_{\tau_m} - Z_{\tau_1} \geq d_m^{(m-1)}(s),\ Z_{\tau_m} \geq d_m^{(m)}(s))
\end{aligned}$$

and

$$d_k^{(m)}(s) = \frac{\log D - s - \left(\delta - \frac{\eta^2}{2}\right)\tau_{m-k}}{\eta}, \quad k = 1, \ldots, m.$$

For $m = 1$ the lemma holds with the convention $x_2^{(1)} = 0$.

We first demonstrate why this lemma proves claim 20. In the lemma above, let $m = n$. Also let the a_i, x_i, d_i be defined as in the previous section. Then the lemma proves claim 20 for the following reasons:

- Since $\tau_i = T - t_{n-i}$ for $i = 1, \ldots, n$, it follows that $x_i^{(n)} = e^{rT} x_i$.
- For $\tau = T$, we have $s = \log S + \left(\delta - \frac{\eta^2}{2}\right) T$, and therefore $d_k^{(n)}(s) = d_k$ for $k = 1,\ \ldots,\ n+1$.
- Finally, the vectors $(W_{t_1},\ \ldots,\ W_{t_{n-1}},\ W_T)$ and $(Z_T - Z_{\tau_{n-1}},\ \ldots,\ Z_T - Z_{\tau_1},\ Z_T)$ have the same distribution, so $a_k^{(n)} = a_k$ for $k = 1,\ \ldots,\ n+1$.

Proof of lemma 23. We prove the lemma by induction on m.

<u>$m = 1$:</u> For $G(s, \tau_1)$ we get

$$\begin{aligned}
G(s, \tau_1) &= E\ G(s + \eta Z_{\tau_1}, 0^+) \\
&= E\left[1 \cdot \mathbb{I}\{s + \eta Z_{\tau_1} \geq \log D\} + e^{-\gamma(1-\rho^2)} \cdot \mathbb{I}\{s + \eta Z_{\tau_1} < \log D\}\right] \\
&= \mathbf{P}\left(Z_{\tau_1} \geq \frac{\log D - s}{\eta}\right) + e^{-\gamma(1-\rho^2)} \cdot \mathbf{P}\left(Z_{\tau_1} < \frac{\log D - s}{\eta}\right) \qquad (6.8) \\
&= a_2^{(1)}(s) e^{\gamma(1-\rho^2) x_2^{(1)}} + a_1^{(1)}(s) e^{\gamma(1-\rho^2) x_1^{(1)}}.
\end{aligned}$$

<u>$m = 2$:</u> We have

$$G(s, \tau_2) = E\ G(s + \eta(Z_{\tau_2} - Z_{\tau_1}), \tau_1^+).$$

The boundary condition at τ_1 gives us

$$G(s, \tau_1^+) = \begin{cases} e^{A\gamma(1-\rho^2)e^{r\tau_1}} \cdot G(s, \tau_1), & s \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_1, \\ e^{-\gamma(1-\rho^2)e^{r\tau_1}}, & s < \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_1. \end{cases}$$

Therefore

$$\begin{aligned} & G(s + \eta(Z_{\tau_2} - Z_{\tau_1}), \tau_1^+) = \\ = \quad & e^{A\gamma(1-\rho^2)e^{r\tau_1}} \cdot G(s + \eta(Z_{\tau_2} - Z_{\tau_1}), \tau_1) \cdot \mathbb{I}\left\{s + \eta(Z_{\tau_2} - Z_{\tau_1}) \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_1\right\} \\ & + e^{-\gamma(1-\rho^2)e^{r\tau_1}} \cdot \mathbb{I}\left\{s + \eta(Z_{\tau_2} - Z_{\tau_1}) < \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_1\right\} \end{aligned} \tag{6.9}$$

Using (6.8) we have

$$\begin{aligned} G(s + \eta(Z_{\tau_2} - Z_{\tau_1}), \tau_1) \quad = \quad & \mathbf{P}\left(Z_{\tau_1} \geq \frac{\log D - s - \eta(Z_{\tau_2} - Z_{\tau_1})}{\eta}\right) \\ & + e^{-\gamma(1-\rho^2)} \cdot \mathbf{P}\left(Z_{\tau_1} < \frac{\log D - s - \eta(Z_{\tau_2} - Z_{\tau_1})}{\eta}\right) \end{aligned} \tag{6.10}$$

The condition $Z_{\tau_1} \geq \frac{\log D - s - \eta(Z_{\tau_2} - Z_{\tau_1})}{\eta}$ is equivalent to $Z_{\tau_2} \geq \frac{\log D - s}{\eta}$, and the condition $s + \eta(Z_{\tau_2} - Z_{\tau_1}) \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_1$ is equivalent to $Z_{\tau_2} - Z_{\tau_1} \geq \frac{\log D - s + \left(\delta - \frac{\eta^2}{2}\right)\tau_1}{\eta}$.

Using (6.10) and writing the probabilities therein as expected values of the corresponding indicator functions, plugging this into (6.9) and finally taking expected values, we get

$$\begin{aligned} G(s, \tau_2) \quad = \quad & E\ G(s + \eta(Z_{\tau_2} - Z_{\tau_1}),\ \tau_1^+) \\ = \quad & e^{-\gamma(1-\rho^2)e^{r\tau_1}} \cdot \mathbf{P}\left(Z_{\tau_2} - Z_{\tau_1} < \frac{\log D - s\left(\delta - \frac{\eta^2}{2}\right)\tau_1}{\eta}\right) \\ & + e^{\gamma(1-\rho^2)[Ae^{r\tau_1} - 1]} \cdot \mathbf{P}\left(Z_{\tau_2} - Z_{\tau_1} \geq \frac{\log D - s\left(\delta - \frac{\eta^2}{2}\right)\tau_1}{\eta},\ Z_{\tau_2} < \frac{\log D - s}{\eta}\right) \\ & + e^{\gamma(1-\rho^2)Ae^{r\tau_1}} \cdot \mathbf{P}\left(Z_{\tau_2} - Z_{\tau_1} \geq \frac{\log D - s\left(\delta - \frac{\eta^2}{2}\right)\tau_1}{\eta},\ Z_{\tau_2} \geq \frac{\log D - s}{\eta}\right) \\ & e^{\gamma(1-\rho^2)x_1^{(2)}} a_1^{(2)}(s) + e^{\gamma(1-\rho^2)x_2^{(2)}} a_2^{(2)}(s) + e^{\gamma(1-\rho^2)x_3^{(2)}} a_3^{(2)}(s), \end{aligned}$$

which is what we claimed.

Induction step $\underline{m \to m+1}$: Suppose that for some $m \in \{2, \ldots, n-1\}$ we have

$$G(s, \tau_m) = \sum_{i=1}^{m+1} a_i^{(m)}(s) e^{\gamma(1-\rho^2)x_i^{(m)}},$$

where the $a_i^{(m)}(s)$, $x_i^{(m)}$ are as in the lemma.

Then $G(s, \tau_{m+1}) = E\, G(s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m}), \tau_m^+)$, where

$$\begin{aligned} G(s, \tau_m^+) &= e^{A\gamma(1-\rho^2)e^{r\tau_m}} G(s, \tau_m) \cdot \mathbb{I}\{s \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_m\} \\ &\quad + e^{-\gamma(1-\rho^2)e^{r\tau_m}} \cdot \mathbb{I}\{s < \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_m\}. \end{aligned}$$

Therefore

$$\begin{aligned} G(s, \tau_{m+1}) &= \sum_{i=1}^{\infty} e^{\gamma(1-\rho^2)\left[Ae^{r\tau_m} + x_i^{(m)}\right]} \cdot E\left[a_i^{(m)}(s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m})) \right. \\ &\quad \left. \cdot \mathbb{I}\{s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m}) \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_m\}\right] \\ &\quad + e^{-\gamma(1-\rho^2)e^{r\tau_m}} \cdot E\left[\mathbb{I}\{s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m}) < \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_m\}\right]. \end{aligned}$$

The term in the last line equals $e^{-\gamma(1-\rho^2)e^{r\tau_m}} \cdot \mathbf{P}\left(Z_{\tau_{m+1}} - Z_{\tau_m} < d_1^{(m+1)}(s)\right)$, which equals $e^{\gamma(1-\rho^2)x_1^{(m+1)}} a_1^{(m+1)}(s)$.

Now we simplify the expected values in the first line. The condition $s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m}) \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_m$ is equivalent to $Z_{\tau_{m+1}} - Z_{\tau_m} \geq d_1^{(m+1)}$. Furthermore,

$$\begin{aligned} a_i^{(m)}(s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m})) &= \mathbf{P}(Z_{\tau_{m+1}} - Z_{\tau_{m-1}} \geq d_2^{(m+1)}(s),\ \ldots, Z_{\tau_{m+1}} - Z_{\tau_{m-i+1}} \geq d_{i-1}^{(m+1)}(s), \\ &\qquad Z_{\tau_{m+1}} - Z_{\tau_{m-i}} < d_i^{(m+1)}(s)). \end{aligned}$$

Hence

$$\begin{aligned} & E\left[a_i^{(m)}(s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m})) \cdot \mathbb{I}\{s + \eta(Z_{\tau_{m+1}} - Z_{\tau_m}) \geq \log D + \left(\delta - \frac{\eta^2}{2}\right)\tau_m\}\right] \\ = \ & \mathbf{P}(Z_{\tau_{m+1}} - Z_{\tau_m} \geq d_1^{(m+1)}(s),\ Z_{\tau_{m+1}} - Z_{\tau_{m-1}} \geq d_2^{(m+1)}(s),\ \ldots, \\ & \qquad Z_{\tau_{m+1}} - Z_{\tau_{m-i+1}} \geq d_{i-1}^{(m+1)}(s),\ Z_{\tau_{m+1}} - Z_{\tau_{m-i}} < d_i^{(m+1)}(s)) \\ = \ & a_{i+1}^{(m+1)}(s). \end{aligned}$$

Finally, $Ae^{r\tau_m} + x_i^{(m)} = x_{i+1}^{(m+1)}$. Putting these together yields

$$G(s, \tau_{m+1}) = \sum_{i=1}^{m+2} a_i^{(m+1)}(s) e^{\gamma(1-\rho^2)x_i^{(m+1)}}.$$

□

Chapter 7

Conclusions and Directions for Future Work

7.1 Conclusions

In this article we introduced a new hybrid model for default occurring in two stages. Firstly, the perceived health of the company, modeled as a GBM, must drop below a critical level leaving the firm in a state of distress – this is the structural part of the model. Once distressed, the firm defaults at an exponential time, viewed as the first arrival of an independent Poisson process – providing the intensity base of the model. The perceived health is not a traded asset, however, it is correlated to the firm's equity and a wide-base (non-defaultable) index. Since the market is incomplete, we utilize certainty equivalence to value credit derivatives written on the firm. When the intensity of the Poisson process driving default in the distressed regime tends to infinity, the barrier for the perceived health behaves as a default barrier and our model reduces to that of Leung, Sircar, and Zariphopoulou (2008). However, in real world settings default will not occur instantly at this point. We succeed in deriving closed form, classical, solutions to the optimization in the absence and presence of the credit risk and hence are able to

determine the certainty equivalent risky yields and CDS spreads.

Given that estimating model parameters from limited data, particularly for the perceived health process, we also develop an uncertain parameter formulation of our model and valuation framework. Motivated by Maenhout (2004) and the robust optimization literature, we introduce a value function which maximizes over admissible trading strategies while minimizing over equivalent measures subject to a scaled entropic penalty. We succeed in obtaining classical solutions to this problem and determine risky yields and CDS spreads subject to parameter uncertainty. All of the observed behaviour is consistent with intuition and we find that parameter uncertainty allows for a wider range of term structures.

We have also begun exploring an extension of the setup above, namely to randomize the boundary below which the firm becomes distressed. This will allow us to introduce a gap in the very term spreads in the healthy regime (recall that a gad appears in the distressed since this regime corresponds to an intensity model). In the complete market setting this has already been addressed, and it is well known that introducing randomness in the default boundary allows the structural model to inherit intensity model features.

7.2 Future Work

There are many doors remaining open for further study, some of which are presented below.

Calibration of parameters. A notoriously difficult issue in structural models is the parameter calibration. In our setup this mainly concerns the parameters of the health process, which one can hope to do by using market data. In the context of this thesis there are two other difficulties. Firstly, one has to determine the risk-aversion parameter γ, which is purely subjective depending on the investor. Secondly, a major problem for the model in chapter 5 is to determine the appropriate penalty term (i.e. the matrix $\mathbf{\Phi}$)

according to the accuracy of the initial model estimation. Maenhout (2004) has made first attempts.

Include jumps in the health process. In practice, short-term credit spreads are non-zero, even if the firm is apparently healthy. A natural way to explain these in a structural framework is to include jumps into the health process C_t. The resulting optimization problems then have to be solved using optimal control for jump diffusions (see e.g Øksendal and Sulem (2007)). It can be expected that numerical methods have to be used for their solution.

Recovery from the distressed regime. In many cases, a firm in financial distress eventually recovers and becomes healthy again. This is a more difficult problem than the one studied here, since now the healthy and distressed regimes will be coupled not only through the boundary condition along the boundary but also through the source terms in the HJB equations. In all, this arena of combining structural and intensity models and incorporating risk-aversion together with parameter uncertainty is a rich area full of interesting and worthwhile problems.

Multiple firms. The main difficulty here is that a high dimensional first passage time problem must be solved. However, if the portfolio has enough symmetry and if the perceived health factors are viewed as uncorrelated the dimensionality reduces considerably. Such an approach, in the purely intensity based model, was explored by Sircar and Zariphopoulou (2009) where the authors demonstrate that the effective correlation can be introduced through risk-aversion alone – without the necessity of correlating the underlying intensity processes.

Different utility functions. While the use of exponential utility is convenient for mathematical tractability, it assumes that the investor has constant absolute risk aversion, which empirically is not true in most cases. As shown in section 1.1, one very unsatisfying property of exponential utility is the fact that in the standard merton investment problem, the optimal investment strategy is a constant (up to discounting), independent

of the investor's current wealth. In the case of power utility $u(x) = x^p$, $0 < p < 1$, or logarithmic utility $u(x) = \log x$, it is not the dollar amount that is constant, but the fraction of total wealth invested in each tradable asset. This result is intuitively more appealing. However, when the investor receives fixed-amount payments, wealth does not separate any more, and consequently one cannot reduce the dimension of the optimization problem at hand. Its solution, as well as the indifference price then has to found numerically. Furthermore, it is not clear, whether there exists a similar substitution as presented in chapter 3 which significantly simplifies the optimization problems at hand.

Bibliography

Anderson, Hansen, and Sargent (2000). Robustness, detection and the price of risk. Working paper, http://www.stanford.edu/∼sargent/research.html.

Bellman, R. (1952). On the theory of dynamic programming. *Proc. Nat. Acad. Sci. USA 38*, 316–319.

Bensoussan, A. (1992). *Stochastic Control with Partial Information*. Cambridge University Press.

Björk, T. (2004). *Arbitrage Theory in Continuous Time*. Oxford Finance.

Black, F. and J. Cox (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance 31*, 351–367.

Brendle, S. and R. Carmona (2004). Hedging in partially observable markets. Technical report, http://www.orfe.princeton.edu/∼rcarmona/fepreprints.html.

Carmona, R. (2008a). Applications to weather derivatives and energy contracts. *Indifference Pricing (ed. Carmona)*, 241–264.

Carmona, R. (2008b). *Indifference Pricing: Theory and Applications*. Princeton University Press.

Cochrane, J. H. (2001). *Asset Pricing*. Princeton University Press.

Davis, M. and A. Norman (1990). Portfolio selection with transaction costs. *Mathematics of Operations Research 15*, 676–713.

Davis, M., V. Panas, and T. Zariphopoulou (1993). *SIAM J. Control and Optimization 31*, 470–493.

Delbaen, F., P. Grandits, T. Rheinländer, D. Sampieri, M. Schweizer, and C. Stricker (2002). Exponential hedging and entropic penalties. *Mathematical Finance 12*, 99–124.

Duffie and Epstein (1992). Stochastic differential utility. *Econometrica 60*(2), 353–394.

Duffie, D. and D. Lando (2001). Term structure of credit spreads with incomplete accounting information. *Econometrica 69*, 633–664.

Fleming, W. and R. Rishel (1975). *Deterministic and Stochastic Optimal Control*. Springer Verlag.

Fleming, W. and H. Soner (2005). *Controlled Markov Processes and Viscosity Solutions*. Springer Verlag.

Föllmer, H. and A. Schied (2002). *Stochastic Finance – An Introduction in Discrete Time*. Walter de Gruyter.

Fouque, J., R. Sircar, and K. Solna (2006). Stochastic volatility effects on defaultable bonds. *Applied Mathematical Finance 13(3)*, 215–244.

Fritelli, M. (2000). The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance 10*(1), 39–52.

Giesecke (2006). Default and information. *Journal of Economic Dynamics and Control 30*, 2281–2303.

Giesecke, K. (2009, March). An overview of credit derivatives. http://creditlab.stanford.edu/Publications.htm.

Henderson, V. and D. Hobson (2002a). Real options with constant relative risk aversion. *Journal of Economic Dynamics and Control 27*, 329–355.

Henderson, V. and D. Hobson (2002b). Substitute hedging. *RISK 15(5)*, 71–75.

Henderson, V. and D. Hobson (2004). Utility indifference pricing – an overview. *Indifference Pricing: Theory and Applications (ed. Carmona)*, 44–73.

Hodges, S. and A. Neuberger (1989). Optimal replication of contingent claims under transaction costs. *Review of Futures Markets 8*, 222–239.

Hull, J. C. (2005). *Options, Futures, and Other Derivatives* (6th ed.). Prentice Hall.

Ilhan, A., M. Jonsson, and R. Sircar (2004). Portfolio optimization with derivatives and indifference pricing. *Indifference Pricing (ed. Carmona)*, 181–210.

Jaimungal, S. and S. Nayak (2004). On valuing equity-linked insurance and reinsurance contracts.

Jaimungal, S. and G. Sigloch (2008). Incorporating risk aversion into structural models of default. Paper submitted for the Bachelier Congress 2008.

Jarrow and Turnbull (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance 50*, 53–85.

Karatzas, I., J. Lehoczky, S. Shreve, and G.-L. Xu (1991). Martingale and duality methods for utility maximization in an incomplete market. *SIAM Journal of Control and Optimization 29*(3), 702–730.

Karatzas, I. and S. Shreve (1998). *Methods of Mathematical Finance*. Springer Verlag.

Kim, I., K. Ramaswamy, and S. Sundaresan (1993). Does default risk in coupons affect the valuation of corporate bonds ? a contingent claims model. *Financial Management 22*, 117–131.

Korn, R. (1997). *Optimal Portfolios*. World Scientific.

Korn, R. (1999). Some applications of impulse controls in mathematical finance. *Mathematical Methods of Operations Research* (50), 493–518.

Kushner, H. (1962). Optimal stochastic control. *IRE Trans. Auto. Control AC-7*.

Leland, H. E. (1977). The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis 12*, 541–552.

Leland, H. E. (1994). Risky debt, bond covenants and optimal capital structure. *Journal of Finance 49*, 1213–1252.

Leland, H. E. and K. B. Toft (1996). Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads. *Journal of Finance 50*, 789–819.

Leung, Sircar, and Zariphopoulou (2008). Credit derivatives and risk aversion. *To be published*.

Lipster, R. and A. Shiryaev (2000). *Statistics of Random Processes, volumes I and II* (2nd ed.). Springer Verlag.

Longstaff, F. and E. Schwartz (1995). A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance 50*, 789–819.

Maenhout (2004). Robust portfolio rules and asset pricing. *The Review of Financial Studies 17*(4).

Merton, R. (1974). On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance 29*, 449–470.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics 51*(3), 247–257.

Øksendal, B. (2007). *Stochastic Differential Equations: An Introduction with Applications* (6th ed.). Springer.

Øksendal, B. and A. Sulem (2007). *Applied Stochastic Control of Jump Diffusions* (2nd ed.). Springer Verlag.

Schönbucher, P. J. (2003). *Credit Derivatives Pricing Models*. Wiley Finance.

Schweizer, M. (1999). A minimality property of the minimal martingale measure. *Statistics and Probability Letters 42(1)*, 27–31.

Sircar, R. and T. Zariphopoulou (2007). *Advances in Mathematical Finance*, Chapter Utility Valuation of Credit Derivatives: Single and Two-Name Cases, pp. 279–301. Birkhauser.

Sircar, R. and T. Zariphopoulou (2009). Utility valuation of credit derivatives and application to cdos. http://www.princeton.edu/∼sircar/, to appear in *Quantitative Finance*.

Stoikov, S. (2005). Princing options from the point of view of a trader. Preprint.

Uppal and Wang (2003). Model misspecification and underdiversification. *The Journal of Finance 58*(6), 2465–2486.

Von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Yong, J. and X. Zhou (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer Verlag.

Young, V. and T. Zariphopoulou (2002). Pricing dynamic insurance risks using the principle of equivalent utility. *Scandinavian Actuarial Journal 4*, 246279.

Zariphopoulou, T. (2001). A solution approach to valuation with unhedgeable risks. *Finance and Stochastics 5(1)*, 61–82.