Optimization of dynamic consumption streams under model uncertainty

vorgelegt von
Diplom-Wirtschaftsmathematikerin
Wiebke Wittmüß
aus Berlin

Von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
Dr.rer.nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. John Sullivan
Berichter/Gutachter: Prof. Dr. rer. nat. Jochen Blath
Berichter/Gutachter: Prof. Dr. rer. nat. Alexander Schied
Berichter/Gutachter: Prof. Dr. oec. Volker Krätschmer

Tag der wissenschaftlichen Aussprache: 28.4.2010

Berlin 2010

D 83

Abstract

The thesis deals with the problem of optimal consumption under uncertainty. Here, we distinguish uncertainty and risk. Risk denotes the fact that the future development of the stock market is random but the investor knows which future scenarios are possible and knows their probabilities - the so-called market measure. In contrast to this we denote by uncertainty the fact that market participants usually do not know the market measure or at least not exactly.

There exists extensive economic literature on the modeling of such a decision problem. In this thesis the agent optimizes his consumption and investment strategy with respect to a robust utility functional.

We investigate the investment problem in a general semimartingale market. The agent can invest an initial capital and a random endowment. To find a solution to the investment problem we use the so-called martingale method. We prove that under appropriate assumptions a unique solution to the investment problem exists and describe it. A further result is that primal and dual problem are convex conjugate functions.

Furthermore we consider a diffusion-jump-model where the coefficients depend on the state of a Markov chain and the investor is uncertain about the intensity of the underlying Poisson process. In this model we consider an agent with logarithmic and HARA utility function. For both we can write the solution of the dual problem as the solution of a system of ordinary differential equations. For this we use stochastic control methods in order to derive so-called Hamilton-Jacobi-Bellmann-equations. The solution to this HJB equation can be determined numerically and we show how thereby the optimal investment strategy can be computed.

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit dem Problem des optimalen Konsums unter Berücksichtigung von Unsicherheit. Dabei unterscheiden wir Unsicherheit und Risiko. Mit Risiko bezeichnen wir die Tatsache, dass die zukünftige Entwicklung des Aktienmarktes zufällig erfolgt, aber gehen davon aus, dass der Investor weiß, welche zukünftigen Szenarien möglich sind und ihre Wahrscheinlichkeit - also das sogenannte Marktmaß - kennt. Im Gegensatz dazu bezeichnen wir mit Unsicherheit die Tatsache, dass Marktteilnehmer normalerweise dieses Marktmaß nicht oder zumindest nicht genau kennen.

Zur Modellierung eines solchen Entscheidungsproblems existiert umfangreiche ökonomische Literatur. In dieser Arbeit optimiert der Agent seine Konsum- und Anlagestrategie bezüglich eines robusten Nutzenfunktionals.

Wir untersuchen das Investitionsproblem in einem allgemeinen Semimartingalmarkt. Der Agent kann ein Anfangskapital und eine zufällige Zusatzausstattung investieren. Um eine Lösung des Investitionsproblems zu finden, wenden wir die sogenannte Martingalmethode an. Wir beweisen, dass unter geeigneten Voraussetzungen eine eindeutige Lösung des Investitionsproblems existiert und beschreiben diese. Ein weiteres Resultat ist, dass primales und duales Problem konjugiert konvexe Funktionen sind.

Weiterhin behandeln wir ein Diffusions-Sprung-Modell, bei dem die Koeffizienten vom Zustand einer Markovkette abhängen und der Investor, sich über die Intensität des zugrunde liegenden Poissonprozesses unsicher ist.In diesem Modell betrachten wir einen Agenten mit logarithmischer und mit HARA Nutzenfunktion. Für beide können wir zunächst die Lösung des dualen Problems als Lösung eines Systems gewöhnlicher Differentialgleichungen darstellen. Dazu nutzen wir stochastische Kontrollmethoden, um sogenannte Hamilton-Jacobi-Bellmann-Gleichungen zu gewinnen. Die Lösung dieser HJB Gleichungen kann numerisch ermittelt werden und wir zeigen, wie man damit die optimale Anlagestrategie berechnen kann.

Acknowledgment

Most importantly, I would like to thank my supervisor Alexander Schied for his guidance, advice, and patience. I am also grateful to Jochen Blath and Volker Krätschmer for writing reports on this thesis.

Special thanks to all my colleagues from TU Berlin Stochastics department for the nice working atmosphere and for the professional as well as non-professional discussions and advice.

Finally, I am really grateful towards my family for their support.

Contents

Chapter 1

Introduction

Ever since the works by Merton [32, 33] in 1969/1971 optimal investment has been one of the major research areas in mathematical finance. Here the interplay between mathematical progress and economic research leads to ever new formulations of the investment problem and possible solutions - starting from the maximization of expected utility several more advanced approaches were developed. One major development was the incorporation of model uncertainty, additional to the consideration of risk. Here risk means that the future market development is random but the randomness behaves according to some known probability model. In contrast to this, model uncertainty or ambiguity is concerned with the fact that the agent does not have a valid probability model for the possible future market development.

In this thesis we consider an agent who invests in a stock market in order to maximize her contentment from consumption and is aware of risk as well as uncertainty. In the first part we investigate the problem of optimal consumption in a general semimartingale framework where the agent may invest in the stock market and receives additional random endowment. We treat the problem with the help of the dual or martingale method. A version of this third chapter was published in [41]. In the second part we use stochastic control techniques to compute the optimal consumption and investment strategy in a specific Markovian switching model.

We specify our problem as a robust optimization problem. In its basic form this specification of the problem is called Maximin Expected Utility (MEU) and was introduced by Gilboa and Schmeidler [18]. One possible interpretation of this approach is that the agent takes a class of possible models into consideration and decides on the worst-case outcome. For the first part of the thesis we use a similar model where the agent applies additionally a penalty function to each possible model. (These utility functionals are closely related to convex risk measures.) Obviously there might as well be agents who are not uncertainty averse but have a more positive attitude towards uncertainty. These preferences can be modeled via α-MEU, which is a mixture of the uncertainty averse and loving attitude via a parameter α. A more detailed explanation of the representation for robust problems is given in the next chapter.

There exists a vast amount of literature for the non-robust problem of optimal terminal wealth discussing different aspects of the problem. For example Kramkov and Schachermayer [28, 29] characterized solutions of this problem in a general setting using duality methods. There exist many extensions of this problem, e.g.

Cvitanić et al. [6], Karatzas and Žitković [27], or Hugonnier and Kramkov [25] consider the case where the agent optimizes the consumption process and also receives some kind of additional random endowment.

In Schied [37], Schied and Wu [40] and Gundel [19] the robust problem (without penalty term) was solved, also via the duality or martingale method.

When the dual method is used one can usually not expect to be able to compute the optimal investment strategy. This is often possible if different methods in more specific models are used: See Duffie and Zariphopoulo [9] for a stochastic control approach for a model with random endowment or Quenez [36] for computation of the optimal strategy in a robust setting. In Becherer [3] and Müller [34] an optimal investment problem is solved using Backward Stochastic Differential Equations. There are several papers on optimal investment in a Markovian switching model, among others compare Bäuerle and Rieder [2]. Also the robust problem with a penalty term was considered, e.g. by Hansen and Sargent [20] and Bordigoni et al. [4] for an entropic penalty term and by Hernández-Hernández and Schied [22, 23] for a general penalty term. For the solution of an optimization problem with α-MEU compare for example Fei [12].

The organization of the thesis is as follows. The second chapter is a short introduction into robust optimization. In the third chapter we use the martingale method to formulate an appropriate dual problem for a robust setting given by a utility function and a penalty term. We prove that the usual duality relations hold and characterize the form of the optimal consumption strategy. In the fourth chapter we apply these results to a Markovian switching model. Here, we do not consider a penalty function or random endowment and work in a rather concrete setting. Using the results of the third chapter we can solve a problem where we need to find the infimum and do not need to compute the saddlepoint of the utility functional. We deal with logarithmic and Hyperbolic Absolute Risk Aversion (HARA) utility. For the case of logarithmic utility we also solve the investment problem for an uncertainty loving investor using *maxmax* expected utility. We do not consider an investor with α-MEU preferences but only the cases $\alpha = 1$ and $\alpha = 0$. Nevertheless, at least for a rather specific setting this seems to be sufficient to solve the investment problem for general α (compare [12]). Furthermore, we give a numerical comparison between an optimistic and a pessimistic investor. Finally, we give an appendix that includes the code that was used to do the simulations and numerical analysis.

Chapter 2

Robust Preferences

One of the major questions when investigating optimization problems is how to model the preferences of an agent. More precisely, we need a numerical representation which enables us to compute which consumption alternative is more attractive to the agent. In this chapter we introduce rather informally the ideas of robust optimization. In the following let $\mathcal{X}$ denote the set of possible payoffs. (Here, we consider just payoffs and no consumption streams. This is sufficient to demonstrate the reasoning behind the robust approach.) These payoffs are subject to future random developments and hence they are random variables on some probability space. We assume that $\mathcal{X}$ contains constants and convex combinations. In our setting the payoffs are the result of investments in the stock-market. If the agent prefers X_2 to X_1 we will write $X_1 \preceq X_2$ and $X_1 \prec X_2$ for strict preference.

It seems rather obvious that agents do not judge investment opportunities just according to their expectation but also consider the likeliness of different outcomes. We call the fact that the payoff is random *risk*. Most agents are risk averse, i.e. if there are two possible payoffs with the same expectation, and one is certain while the other is random, most agents prefer the certain payoff.

The above definition of risk aversion uses the *expectation* of the payoff. It seems sensible that this expectation should be taken with respect to the market measure. However, there are two problems that will be discussed in more detail below. Firstly, even in situations where the probability model for the payoff is given, the agent might use a *subjective* probability to evaluate the situation. Secondly, it is not realistic that any market model reflects market developments correctly and also given a class of possible models the agent might feel *uncertain* about the correct market measure. The theory of robust preferences tries to find a model that allows to incorporate this *uncertainty* or *ambiguity*.

The theory of preferences can be applied to a much wider range of consumption than just monetary goods but this is sufficient for our purpose. All of the numerical representations below can be characterized by a set of axioms.

The following set of axioms can be used to characterize the *subjective expected utility* as introduced by Savage.

(i) (Completeness) For all $X_1, X_2 \in \mathcal{X}$ we have either $X_1 \prec X_2$, $X_2 \prec X_1$ or $X_1 \approx X_2$. I.e., the agent faced with two alternatives knows which alternative she prefers or if both are equally desirable.

(ii) (Transitivity) For all $X_1, X_2, X_3 \in \mathcal{X}$ with $X_1 \preceq X_2$ and $X_2 \preceq X_3$ we have $X_1 \preceq X_3$.

(iii) (Continuity) If $X_1 \prec X_2 \prec X_3$ then there exists some $\alpha \in [0, 1]$ such that $\alpha X_1 + (1 - \alpha)X_3 \prec X_2$.

(iv) (Independence) If $X_1 \preceq X_2$, $X_3 \in \mathcal{X}$ and $\alpha \in [0, 1]$ then

$$\alpha X_1 + (1 - \alpha)X_3 \preceq \alpha X_2 + (1 - \alpha)X_3.$$

This means that if X_2 is preferred to X_1 this preference is not changed if another payoff is added.

(v) (Monotonicity) If $X_1(\omega) \leq X_2(\omega)$ for all $\omega \in \Omega$ then $X_1 \preceq X_2$. I.e. if the payoff X_2 is almost surely higher than X_1 the agent will prefer X_2.

(vi) (Non-degeneracy) There exist X_1 and $X_2 \in \mathcal{X}$ such that $X_1 \prec X_2$.

These axioms guarantee a numerical representation of the form

$$E_Q[\,U(X)\,].$$

Hence, the representation is given as a von Neumann-Morgenstern expected utility with respect to the subjective probability measure Q which is typically different from the market measure and not the same for all agents. Here, U is a utility function.

We will assume the investor to be risk averse. Apart from technical issues the utility function has then two main features: it is concave and increasing. The monotonicity is implied by Axiom (v) and the concavity corresponds to the risk aversion. In chapter 4 we will work with Hyperbolic absolute risk aversion (HARA) utility functions. These include logarithmic utility, i.e. $U(x) = \log(x)$ and for a risk aversion parameter $\alpha \neq 0$, $\alpha < 1$ utility functions of the form $U(x) = x^{\alpha}/\alpha$.

Now we will concentrate on the above axioms and the question how they can be changed in order to include uncertainty. Axioms (i), (ii) and (v) are obviously necessary and should be satisfied by most agents. Axiom (iii) and (vi) are of technical importance. Axiom (iv) is the axiom that is often violated by agents faced with ambiguity and which will therefore be the main focus of the following consideration.

Ellsberg [11] showed by the following experiment that the independence axiom does not reflect reality very well in situations where uncertainty/ambiguity is present:

Suppose there is an urn containing 30 red balls and 60 balls that are either black or yellow. The agent can choose between bet A and B. In case of bet A he wins \$ 100 if a red ball is drawn (otherwise nothing) in case of bet B he receives \$ 100 if a yellow ball is drawn. In this situation most people prefer bet A. Now the bets are changed: In both cases the agent also wins if a black ball is drawn. Now most agents decide for the second bet. This is a contradiction to the independence axiom and these kind of preferences cannot be modeled via the Savage representation. Ellsberg explained this contradiction to the independence axiom by ambiguity - the agent perceives/has a lack of information about the probability of drawing a favorable ball. Agents who decide for bet A in the first situation are typically uncertainty averse and assume the unknown probability of drawing a yellow ball as small. In the second situation the black balls work as a hedge giving a 60:100 chance to win for bet B.

Gilboa and Schmeidler [18] developed a set of axioms where they replaced the independence axiom by a *certainty independence axiom*, i.e.

(iv') $X_1 \prec X_2$ if and only if $\alpha X_1 + (1-\alpha)x \prec \alpha X_2 + (1-\alpha)x$ for all constants x and $\alpha \in (0,1)$.

This means a *certain* payoff does not change the preferences of the investor. This certain payoff does not admit hedging and hence the situation of the Ellsberg paradox is avoided. Furthermore, Gilboa and Schmeidler assume Axiom (i)- (iii), (v), (vi) and additionally

(vii) (Uncertainty aversion) For $X_1, X_2 \in \mathcal{X}$ with $X_1 \approx X_2$ and all $\alpha \in [0,1]$ we have $\alpha X_1 + (1-\alpha)X_2 \succeq X_1$.

This axiom states that the investor will always prefer a possible hedge.

Here, the Ellsberg paradox can be resolved (at least for uncertainty averse agents). The preferences given by these axioms have a representation via

$$\inf_{Q \in \mathcal{Q}} E_Q[\,U(X)\,] \tag{2.1}$$

where $\mathcal{Q}$ is a closed convex set of probability measures.

One possible interpretation of this approach is that the agent takes a class of possible models into consideration and decides on the worst-case outcome.

This interpretation also explains why it is useful to extend the robust approach to an evaluation that includes a penalty function: the agent will typically have some assessment of the different models (for instance she might use estimated data for the model specification, and while she does not assume her estimates to be correct, she believes values close to the estimates to be more likely than values far off). It is possible to incorporate this evaluation by the introduction of a penalty function, which allows to adjust the impact of a model according to its plausibility.

In [31] Maccheroni et al. weaken the axiom of *certainty independence* to

(iv") (Weak certainty independence) For all $X_1, X_2 \in \mathcal{X}$, constants $x, y \in \mathcal{X}$ and $\alpha \in (0,1)$ we have

$$\alpha X_1 + (1-\alpha)x \prec \alpha X_2 + (1-\alpha)x \Rightarrow \alpha X_1 + (1-\alpha)y \prec \alpha X_2 + (1-\alpha)y.$$

Axiom (iv') could also be written as

For all $X_1, X_2 \in \mathcal{X}$, constants $x, y \in \mathcal{X}$ and $\alpha, \beta \in (0,1)$ we have

$$\alpha X_1 + (1-\alpha)x \prec \alpha X_2 + (1-\alpha)x \Rightarrow \alpha X_1 + (1-\beta)y \prec \alpha X_2 + (1-\beta)y.$$

In this form it is easier to compare both axioms: The choice of α and β influences how close the combined payoff is to a constant. Axiom (iv') states that the preferences of the agent do not depend on the ratio of certain to uncertain payoff. In contrast to this, Axiom (iv") means that the agent may have preferences depending on this ratio.

A preference order satisfying these axioms has a representation via

$$\inf_{Q\in\mathcal{Q}}(E_Q[\,U(X)\,]+\gamma(Q))$$

where γ is such a penalty function. We will use this representation of preferences in Chapter 3. Earlier e.g. Hansen and Sargent [20] developed a similar model where they use relative entropy as a penalty function.

In this thesis we will also use the link between the robust preference functionals and *monetary risk measures.* A risk measure assigns to a random loss/gain the amount of money the agent needs to add in order to reduce the risk to an "acceptable" level. A detailed discussion of these risk measures can be found in [15, Chapter 4]. If $\rho : L^\infty \to \mathbb{R}$ satisfies the first two axioms of the following set of axioms it is called a monetary risk measure. If it satisfies (RM1)-(RM3) it is called a convex risk measure, and (RM1)-(RM4) characterize a coherent risk measure.

(RM1) (Monotonicity) If $X_1(\omega) \leq X_2(\omega)$ for all $\omega \in \Omega$ then $\rho(X_1) \geq \rho(X_2)$.

(RM2) (Cash invariance) For a constant x we have $\rho(X + c) = \rho(X) - c$.

(RM3) (Convexity) For $\alpha \in (0, 1)$ we have

$$\rho(\alpha X_1 + (1-\alpha)X_2) \leq \alpha\rho(X_1) + (1-\alpha)\rho(X_2).$$

(RM4) (Positive homogeneity) If $\alpha \geq 0$ then $\rho(\alpha X) = \alpha\rho(X)$.

Axiom (i) and (ii) are clear in view of the interpretation of $\rho(X)$ as capital requirement. The convexity implies that a diversified portfolio has a smaller risk and hence, encourages hedging.

It is now interesting that these axioms yield a representation similar to (2.1). Namely, for coherent risk measures we have

$$\rho(X) = \sup_{Q\in\mathcal{Q}} E_Q[-X].$$

The more general notion of a convex risk measure has a representation via

$$\rho(X) = \sup_{Q\in\mathcal{Q}}(E_Q[-X] + \gamma(Q)).$$

This analogous representation will be used in the thesis.

Another aspect when considering ambiguous situations is the attitude of the agent towards uncertainty. In the above approaches the agent is always assumed to be uncertainty averse as given by Axiom (vii). This means the agent prefers always a possible hedge. However, there might be agents that have a more distinguished attitude towards uncertainty. The most trivial case would be the preferences of an uncertainty loving investor, these can be modeled via $\sup_{Q\in\mathcal{Q}} E_Q[\,U(X)\,]$. We work with this kind of preferences in the fourth chapter where we call these agents optimists.

In order to make it possible to distinguish between perceived uncertainty (i.e. the set $\mathcal{Q}$) and the attitude towards this uncertainty, Ghirardato et al. [17] work with the axioms (i)-(iii), (iv'), (v) and (vi) given by Gilboa and Schmeidler [18] and change Axiom (vii) of uncertainty aversion. They show that the resulting preferences have a numerical representation and that α-MEU is a subclass for which Axiom (vii) is replaced by their Axiom 7. Here, α-MEU is the representation we get if we take a convex mixture of the "classical" and the "optimistic" MEU: the agent evaluates the payoff X via

$$\alpha \inf_{Q\in\mathcal{Q}} E_Q[\,U(X)\,] + (1-\alpha) \sup_{Q\in\mathcal{Q}} E_Q[\,U(X)\,] \text{ for } \alpha \in [0,1].$$

We will only consider the cases $\alpha = 0$ and $\alpha = 1$. Nevertheless, it would be interesting to investigate also more general cases.

Chapter 3

Duality theory

This chapter is based on [41]. We will investigate the problem of optimal consumption in a general semimartingale framework where the agent may invest in the stock market and receives additional random endowment. She evaluates positions by a robust utility functional described through a utility function U and a penalty term γ. More specifically the agent tries to maximize

$$\inf_{Q \ll P} \left(E_Q\Big[\int_0^T U(t, c_t)\, \mu(dt) \Big] + \gamma(Q) \right)$$

over all possible consumption rate processes c. First the dual function for this problem is given. As usual in the theory of utility maximization the dual problem is a minimization problem where the infimum is taken over the dual set, that is somehow related to the set of equivalent martingale measures. In our case the dual solution will consist of two elements - $(\widehat{Q}, \widehat{R})$. Here, $\widehat{Q} \ll P$ is the "worst-case scenario" and $\widehat{R}$ is the minimizer from the dual set. We examine properties and relations between primal and dual problem, e.g. we will show that both problems have a solution and are conjugate to one another. Next we will verify that the above maximin problem and the corresponding minimax problem are equivalent.

The problem is motivated by [38] where Schied considers the problem of optimal terminal wealth with respect to a convex risk measure in a setting without random endowment. We extend this problem by introducing a concept of consumption, general enough to include also the maximization of terminal wealth and additional random endowment (compare Karatzas and Žitković [27]). In contrast to the problem of optimal terminal wealth, the problem of optimal consumption does not admit an equivalent static version. Hence, we cannot restrict ourselves to random variables to solve the primal problem, but need to work with stochastic processes. Furthermore, our dual problem is different from the dual problem in Schied [38] or Karatzas and Žitković [27] due to random endowment on the one hand and robustification on the other. These differences imply for instance that there are cases where, given the dual solution $(\widehat{Q}, \widehat{R})$, $\widehat{R}$ is no longer a solution to the associated dual problem under the model $\widehat{Q}$. Aspects regarding the dual function will be discussed in more detail in the third section. Robustification changes the problem of optimal consumption also significantly with respect to the optimization procedure since we consider a maximin instead of just a maximization problem. This extension yields some new results and partly we give new proofs for known results.

In [5] Burgert and Rüschendorf also consider a robust version of [27]. The most obvious difference to our setting is that we deal with preferences given by convex instead of just coherent risk measures. Furthermore, Burgert and Rüschendorf work under the serious restriction that the set $\mathcal{Q}$ should only consist of measures that are equivalent to P and have a uniformly bounded density. The first assumption rules out risk measures such as AVaR, the second standard dynamic consistent coherent utility functionals in a Brownian setting such as used in [22].

In the next section we will introduce the market setting and describe the agent's preferences in more detail before we state our main theorem and give an example. In the third section we will derive our dual problem and prove the statements of the theorem related to it. Finally, in the fourth section we will finish the proof of the theorem.

3.1 Notation and main results

We consider an agent who wants to maximize her utility from consumption between time zero and some finite time horizon T. She is endowed with an initial capital and receives additional random endowment over time, which she may invest into d assets. To formalize this problem we use the same market model as Karatzas and Žitković [27]. That means we model the price process of the assets as a d-dimensional RCLL semimartingale on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, P)$ where the filtration satisfies the usual conditions. The financial market is assumed to be arbitrage-free in the sense that the set $\mathcal{M}$ of supermartingale measures equivalent to P is not empty. The portfolio process is denoted by $\theta = (\theta_t)_{0\le t\le T}$, and we allow only those strategies for which $\int_0^t \theta_u \, dS_u$ is bounded from below by some constant.[1] The initial capital is denoted by x and the random endowment is described as a non-decreasing, adapted, RCLL process $\mathcal{E} = (\mathcal{E}_t)_{0\le t\le T}$ with $\mathcal{E}_T \in L^\infty(P)$. The consumption process $C = (C_t)_{0\le t\le T}$ is assumed to be a non-negative, non-decreasing and adapted RCLL process. More specifically we will only deal with consumption processes that can be written as

$$C_t = \int_0^t c_s \, \mu(ds)$$

where μ is a probability measure on $[0,1]$ which is diffuse on $[0,1)$. This implies that the agent consumes in a continuous way except for the final time T where we also allow for lump consumption. In particular we can choose $\mu = \delta_{\{T\}}$ which corresponds to the problem of optimizing terminal wealth.

The terminal wealth is required to be non-negative:

$$x + \mathcal{E}_T + \int_0^T \theta_t \, dS_t - C_T \ge 0 \qquad P-\text{a.s.} \tag{3.1}$$

We denote by $\mathcal{A}(x,\mu)$ the set of consumption process densities $(c_t)_{0\le t\le T}$, for which an admissible strategy θ exists, such that condition (3.1) is fulfilled.

[1] This condition will correspond to our admissibility condition for strategies, since the additional income of the agent is bounded.

The utility from consumption is measured in terms of a robust utility functional of the form

$$c \longmapsto \inf_{Q \ll P} \Big(E_Q\Big[\int_0^T U(t, c_t)\, \mu(dt) \Big] + \gamma(Q) \Big). \tag{3.2}$$

This functional was introduced in the second chapter and is closely related to the convex risk measure

$$\rho(Y) := \sup_{Q \ll P} \big(E_Q[-Y] - \gamma(Q) \big), \qquad Y \in L^\infty(P). \tag{3.3}$$

More precisely the penalty function γ is supposed to be bounded from below and equal to the minimal penalty function of the convex risk measure above. In other words, it satisfies the biduality relation

$$\gamma(Q) = \sup_{Y \in L^\infty(P)} \big(E_Q[-Y] - \rho(Y) \big). \tag{3.4}$$

Furthermore, we need the following conditions on γ; compare [38, Assumption 1].

Assumption 3.1
We assume that the risk measure ρ is continuous from below, i.e. for a sequence $(Y_n) \subset L^\infty(P)$ increasing a.s. to some $Y \in L^\infty(P)$, we have $\rho(Y_n) \searrow \rho(Y)$. Furthermore, ρ needs to be sensitive in the sense that $\rho(Y)$ is strictly positive for all $Y \in L^\infty_-(P) \setminus \{0\}$.

We work with a utility function $U\colon [0,T] \times \mathbb{R}^+ \to \mathbb{R}$ with the properties assumed in Definition 3.1 of [27]:

Assumption 3.2
For fixed $t \in [0,T]$ we request $U(t,.)\colon \mathbb{R}^+ \to \mathbb{R}$ to be a utility function, i.e. $U(t,.)$ is strictly concave, increasing, continuously differentiable and satisfies the so-called Inada conditions, namely

$$\lim_{x \to 0} U_x(t,x) = \infty \quad \text{and} \quad \lim_{x \to \infty} U_x(t,x) = 0$$

for all $t \geq 0$.

The marginal utility is bounded by the strictly decreasing continuous functions K_1 and K_2, such that

$$K_1(x) \leq U_x(t,x) \leq K_2(x)$$

and

$$\limsup_{x \to \infty} \frac{K_2(x)}{K_1(x)} < \infty.$$

The map $t \mapsto U(t,1)$ is bounded and

$$\lim_{x \to \infty} \inf_{t \in [0,T]} U(t,x) > 0.$$

Additionally, U is supposed to be of reasonable asymptotic elasticity, i.e.

$$\limsup_{x \to \infty} \Big(\sup_t \frac{x\, U_x(t,x)}{U(t,x)} \Big) < 1.$$

See [27] for a discussion of these assumptions.

To avoid problems of evaluating

$$E_Q\Big[\int_0^T U(t,c_t)\,\mu(dt)\Big] + \gamma(Q)$$

for the case where U is unbounded from below it is sensible to restrict the set of measures that enter the optimization problem to

$$\mathcal{Q} = \{Q \ll P \,|\, \gamma(Q) < \infty\}.$$

Hence, our optimization problem is now

$$\text{maximize } \inf_{Q\in\mathcal{Q}}\bigg(E_Q\Big[\int_0^T U(t,c_t)\,\mu(dt)\Big] + \gamma(Q)\bigg) \text{ over all } c \in \mathcal{A}(x,\mu).$$

Furthermore, to circumvent difficulties when integrating we follow [38] in setting

$$E_Q[\,F\,] := \sup_n E_Q[\,F \wedge n\,] = \lim_n E_Q[\,F \wedge n\,] \qquad \text{for arbitrary } F \in L^0. \tag{3.5}$$

In doing so we keep the functional

$$(c_t)_{0\le t\le T} \longmapsto E_Q\Big[\int_0^T U(t,c_t)\,\mu(dt)\Big]$$

concave.

In the following we will use $\mathcal{U}_Q(c)$ to abbreviate $E_Q\Big[\int_0^T U(t,c_t)\,\mu(dt)\Big]$. We define the value function of the maximization problem as follows

$$u(x) = \sup_{c\in\mathcal{A}(x,\mu)} \inf_{Q\in\mathcal{Q}}\big(\mathcal{U}_Q(c) + \gamma(Q)\big). \tag{3.6}$$

Hence, the investor needs to solve a maximin problem in order to find u. Another approach to the optimization problem is to solve the problem in each possible model, i.e. compute

$$u_Q(x) = \sup_{c\in\mathcal{A}(x,\mu)} \mathcal{U}_Q(c)$$

for each $Q \in \mathcal{Q}$ and then robustify the result by taking the infimum. We will show that both methods lead to the same result, namely

$$u(x) = \inf_{Q\in\mathcal{Q}}(u_Q(x) + \gamma(Q)).$$

Sometimes it will be more convenient to work with densities instead of measures. We denote the density of Q with respect to P on $\mathcal{F}_T$ by $Z^Q = dQ/dP$ and the set $\{Z^Q \,|\, Q \in \mathcal{Q}\}$ by $\mathcal{Z}$. We will identify Q and Z^Q, and thus $\gamma(Q)$ and $\gamma(Z^Q)$ or u_Q and u_{Z^Q} denote the same object. While we write Z^Q for $dQ/dP|_{\mathcal{F}_T}$ we will denote the corresponding density process by $(Z_t^Q)_{0\le t\le T}$.

In this chapter we will use the dual (or martingale) approach to characterize the solution to our optimization problem. A downside of the dual method is that it does not directly give the investment strategy θ which is necessary in order to realize the optimal consumption plan. In the next chapter we will give an example how to compute this strategy for a special market model. We will do so by applying stochastic control techniques to the dual problem. The papers [22] and [36] are further examples where the optimal strategy is determined for special cases. These authors also argue that the dual problem is easier to treat than the primal problem which shows that our results are still useful. Furthermore, in this chapter we do not restrict the form of the penalty function, therefore our approach covers also the cases where time consistency is lacking and hence optimal control techniques cannot be applied (see [38]).

In general the dual problem is an associated minimization problem where the dual domain is related to the set of equivalent martingale measures. As was shown in [6] and [27] we need to use $\mathcal{D}$, the weak*-closure of the set of equivalent supermartingale measures $\mathcal{M}$, as dual domain. More precisely, we identify the set $\mathcal{M}$ with its embedding in the dual of $L^\infty(P)$, $(L^\infty(P))^*$. Then $\mathcal{D}$ is the $\sigma((L^\infty(P))^*, L^\infty(P))$-closure of $\mathcal{M}$. The set $\mathcal{D}$ contains also finitely-additive measures to which we cannot directly associate a density process. Therefore, we use that each R in $\mathcal{D}$ has a unique Hewitt-Yosida decomposition $R^r + R^s$ where the regular part R^r is the maximal countable measure on $\mathcal{F}$ that is dominated by R. Hence, we can for each $R \in \mathcal{D}$ define a supermartingale L^R, where L^R_t is the density of the regular part $(R|_{\mathcal{F}_t})^r$ of $R|_{\mathcal{F}_t}$ with respect to $P|_{\mathcal{F}_t}$. In the following we will work with the RCLL supermartingale Y^R that coincides with L^R for all $t \in \mathbb{Q} \cap [0,T]$. For a proof of the existence of Y^R and further properties of Y and $\mathcal{D}$ see [27]. We need as further notation $\langle R, \mathcal{E}_T\rangle$ which gives the canonical pairing. Observe that in particular $\langle R, I_\Omega\rangle = 1$.

Using $\mathcal{D}$ we consider as dual problem

$$v(y) = \inf_{Z\in\mathcal{Z}} \left(\inf_{R\in\mathcal{D}} \Big(E\Big[Z \int_0^T V\Big(t, y\frac{Y_t^R}{Z_t}\Big)\mu(dt)\Big] + y\langle R, \mathcal{E}_T\rangle \Big) + \gamma(Z) \right) \tag{3.7}$$

where V is the convex conjugate of U, i.e.

$$V(t,y) = \sup_{x\geq 0}(U(t,x) - xy).$$

In section 3.2 we will explain how this problem can be derived from the capital constraint (3.1). To simplify notation, $E_Q\Big[\int_0^T V(t, Y_t^R)\,\mu(dt)\Big]$ will be denoted by $\mathcal{V}_Q(Y^R)$. We defined the function u_Q as the solution to the optimization problem under the subjective probability Q. We can now define an associated dual value function v_Q as

$$v_Q(y) = \inf_{R\in\mathcal{D}} \big(\mathcal{V}_Q(yY^R/Z^Q) + y\langle R, \mathcal{E}_T\rangle \big).$$

Remark 3.3
In the introduction convex risk measures were interpreted as a worst-case approach for different scenarios Q. Using this interpretation it seems sensible that when the

agent evaluates the utility in a scenario given by a measure Q, all strategies that satisfy the capital constraint under this measure should be admissible. (And hence, the dual domain would depend on the measure Q.) But we formulated the assumptions for the market and the trading strategy only with respect to the measure P. One reason why we use P is that we use the convex risk measure as a model for the agent's preferences, not as a model of the "real" market. Furthermore, the worst-case measure may allow for arbitrage, compare [38, Example 3.2]. Hence, to exclude arbitrage opportunities it is necessary to formulate the admissibility condition under P.

Nevertheless, for all Q out of

$$\mathcal{Q}_e := \{Q \in \mathcal{Q} \,|\, Q \sim P\}$$

we have the conditions necessary to apply standard duality results. (Lemma 3.11 will guarantee that $\mathcal{Q}_e$ is nonempty.)

Assumption 3.4
In the following we will assume that there exists $Q_0 \in \mathcal{Q}_e$ that satisfies

$$u_{Q_0}(x) < \infty \text{ for some } x > 0.$$

This is a similar assumption as is needed in [27] to guarantee the existence of solutions to both the primal and the dual problems under the subjective probability measure Q_0. Furthermore, we can conclude that u_{Q_0} and v_{Q_0} are dual functions [27, Theorem 3.10]. We will show that we have similar results in our robust setting.

Theorem 3.5
Under the above assumption the following assertions are valid.

1. *Both value functions u and v take only finite values and satisfy*

$$u'(\infty-) = 0 \qquad \text{and} \qquad v'(0+) = -\infty,$$

u is strictly concave and v is continuously differentiable.

2. *The value function u satisfies*

$$u(x) = \sup_{c \in \mathcal{A}(x,\mu)} \inf_{Q\in\mathcal{Q}} \big(\mathcal{U}_Q(c) + \gamma(Q)\big) = \inf_{Q\in\mathcal{Q}} \Big(\sup_{c \in \mathcal{A}(x,\mu)} \mathcal{U}_Q(c) + \gamma(Q)\Big).$$

3. *The two value functions u and v are conjugate to each other:*

$$u(x) = \inf_{y>0} \big(v(y) + xy\big) \qquad \text{and} \qquad v(y) = \sup_{x>0} \big(u(x) - xy\big).$$

In particular, v is convex.

4. *The derivative of v satisfies*

$$v'(\infty-) \in [\, \inf_{R\in\mathcal{D}} \langle R, \mathcal{E}_T\rangle, \sup_{R\in\mathcal{D}} \langle R, \mathcal{E}_T\rangle \,]. \tag{3.8}$$

If $\mathcal{E}_T \equiv 0$ the derivatives of v and u satisfy

$$u'(0+) = \infty \qquad \text{and} \qquad v'(\infty-) = 0. \tag{3.9}$$

5. *There exists a solution $(\widehat{Q}, \widehat{R}) \in \mathcal{Q} \times \mathcal{D}$ to the dual problem, i.e.*

$$v(y) = \mathcal{V}_{\widehat{Q},\mu}\Big(y\,\frac{Y^{\widehat{R}}}{Z^{\widehat{Q}}}\Big) + y\,\langle \widehat{R}, \mathcal{E}_T\rangle + \gamma(\widehat{Q}).$$

6. *For any $x > 0$ there exists an optimal consumption strategy $\hat{c} \in \mathcal{A}(x,\mu)$. If $(\widehat{Q}, \widehat{R})$ is a solution to the dual problem for $y > 0$ such that $x = -v'(y)$ then*

$$u(x) = \inf_{Q\in\mathcal{Q}}\Big(\mathcal{U}_Q(\hat{c}) + \gamma(Q)\Big) = \mathcal{U}_{\widehat{Q},\mu}(\hat{c}) + \gamma(\widehat{Q}) = u_{\widehat{Q}}(x) + \gamma(\widehat{Q})$$

and

$$\hat{c} = I\Big(\cdot, \frac{\widehat{y}\,Y^{\widehat{R}}}{\widehat{Z}}\Big) \qquad \widehat{Q}\otimes\mu\text{-a.s.},$$

where

$$\widehat{Z} = \frac{d\widehat{Q}}{dP} \quad \text{and} \quad I(t,\cdot) = \Big(\frac{\partial U}{\partial x}(t,\cdot)\Big)^{-1}.$$

Obviously the problem is easier to treat if $\mathcal{Q}$ contains only measures that are equivalent to P. As in Schied [38] we get additional results in this case:

Corollary 3.6
If Assumption 3.4 and $\mathcal{Q} = \mathcal{Q}_e$ are satisfied and γ is strictly convex on $\mathcal{Q}$ then the value function u is continuously differentiable, the dual value function v is strictly convex, and for each $y > 0$ there exist $\widehat{Q} \in \mathcal{Q}$ and $\widehat{R} \in \mathcal{D}$ such that

$$v(y) = \mathcal{V}_{\widehat{Q},\mu}\Big(y\,\frac{Y^{\widehat{R}}}{Z^{\widehat{Q}}}\Big) + y\,\langle \widehat{R}, \mathcal{E}_T\rangle + \gamma(\widehat{Q}).$$

Moreover, $Y^{\widehat{R}}$ is unique. For any $x > 0$, the optimal solution $\hat{c} \in \mathcal{A}(x,\mu)$ is P-a.s. unique.

The condition of strict convexity for γ is crucial for the differentiability of u.

Remark 3.7
In the case $\mathcal{E} = 0$ the finitely additive part of $R \in \mathcal{D}$ does not occur in the formulation of the dual problem. We can thus replace the set $\mathcal{D}$ by the set

$$\mathcal{Y}^{\mathcal{D}} = \{Y^R \,|\, R \in \mathcal{D}\}.$$

In fact, one can easily check that Theorem 2.10 in [27] implies that $\mathcal{Y}^{\mathcal{D}}$ can be replaced by the even more convenient set $\mathcal{Y}(1)$, which is as usual defined as the set of all non-negative supermartingales starting in one for which YX is a supermartingale for all admissible value processes X, also starting in one.

The following example considers the case of uncertain drift in a Black-Scholes model. Here one can see the advantage of a penalty function - if we used an MEU approach the least absolute drift would always give the optimum but when a penalty is introduced the result is dependent on the drift under the measure P.

This is obviously an extremely easy example. In the next chapter we will give an example where we apply our results to a more complex market model. (Compare also e.g. [39] or [24]).

Example 3.8
We consider a Black-Scholes model, more precisely the dynamics of S and the money market account B are given through

$$\begin{aligned} dB_t =& B_t r_t dt, & B_0 =& 1, \\ dS_t =& S_t(b_t dt + \sigma dW_t), & S_0 =& s_0 > 0 \end{aligned}$$

for constant $\sigma > 0$, progressively measurable b and r and a Brownian motion W. This market is complete. Hence, the set of martingale measures consists only of one element and we do not need to consider different processes Y^R, $R \in \mathcal{D}$ but need only

$$Y_t = \exp\Big(-\int_0^t \frac{b_s - r_s}{\sigma}\, dW_s - \frac{1}{2}\int_0^t \frac{(b_s - r_s)^2}{\sigma^2}\, ds\Big).$$

It is well-known that for each measure $Q \ll P$ there exists a progressively measurable η such that the density dQ/dP can be written as

$$\frac{dQ}{dP} = \exp\Big(\int_0^T \eta_s\, dW_s - \frac{1}{2}\int_0^T \eta_s^2\, ds\Big) \qquad Q - \text{a.s.};$$

see, e.g., [21, Lemma 3.1].

As penalty function we take the relative entropy between Q and P for Q "similar" to P, i.e.

$$\gamma(Q) = \begin{cases} E_Q[\int_0^T \eta_s\, dW_s - \frac{1}{2}\int_0^T \eta_s^2\, ds], & \text{if } |\eta| \le 1 \\ \infty, & \text{otherwise.} \end{cases}$$

and we choose the non-time-dependent utility functional $U(t,x) = \log(x)$.

The convex conjugate of U is given by $V(t,y) = -1 - \log(y)$.

We set

$$\tilde{b}_s = \sigma\eta_s + b_s$$

which corresponds to the drift-rate of S under Q. We will denote the Q-Brownian motion

$$W - \int (\tilde{b}_s - b_s)/\sigma\, ds$$

by W^Q. The condition $|\eta| \le 1$ translates to $\tilde{b} \in [b - \sigma, b + \sigma]$.

With this we have

$$
\begin{aligned}
v(y) &= \inf_{z\in\mathcal{Z}}\Big(E\big[Z\int_0^T V(t, yY_t/Z_t)\,\mu(dt)\big] + \gamma(Z)\Big) \\
&= \inf_{\tilde b\in[-\sigma+b,\sigma+b]}\Bigg(E_Q\Big[-\int_0^T \frac{b_t-\tilde b_t}{\sigma}\,dW_t^Q + \frac{1}{2}\int_0^T \frac{(b_t-\tilde b_t)^2}{\sigma^2}\,dt\Big] \\
&\qquad + E_Q\Big[\int_0^T\Big(-1-\log y + \int_0^t \frac{\tilde b_s - r_s}{\sigma}\,dW_s^Q + \frac{1}{2}\int_0^t \frac{(\tilde b_s - r_s)^2}{\sigma^2}\,ds\Big)\,\mu(dt)\Big]\Bigg).
\end{aligned}
$$

If we consider the case where $\mu = \delta_{\{T\}}$ we get

$$
v(y) = \inf_{\tilde b\in[-\sigma+b,\sigma+b]}\Bigg(\log 1/y - 1 + E_Q\Big[\frac{1}{2}\int_0^T \frac{(\tilde b_t - r_t)^2}{\sigma^2}\,dt + \frac{1}{2}\int_0^T \frac{(b_t-\tilde b_t)^2}{\sigma^2}\,dt\Big]\Bigg).
$$

If we minimize

$$
\frac{(\tilde b_t - r_t)^2}{\sigma^2} + \frac{(b_t-\tilde b_t)^2}{\sigma^2}
$$

with respect to $\tilde b$ we get a unique minimum at $(b_t + r_t)/2$. Obviously this might fail to satisfy the restrictions which implies that the optimum is at a boundary. Together this yields that the above infimum is achieved for

$$
\tilde b_t^* = \begin{cases} b_t - \sigma, & \text{if } (r_t+b_t)/2 < b_t - \sigma \\ (b_t + r_t)/2, & \text{if } (r_t+b_t)/2 \in [b_t-\sigma, b_t+\sigma] \\ b_t+\sigma, & \text{if } (r_t+b_t)/2 > b_t+\sigma. \end{cases}
$$

First we observe, that the $\tilde b^*$ depends on b which shows the influence of the penalty function. Since σ is typically rather small $\tilde b^*$ will often be at the boundary if b and r are not close. This results from our choice of the restriction of η. Obviously we could change this to $|\eta| \le c$ for some constant $c > 0$.

Furthermore if we assume that $\tilde b^* = (b+r)/2$, we can compute via $u(x) = \inf_{y>0}(v(y) + xy)$ that

$$
u(x) = \log x + E_Q\Big[\int_0^T \frac{(b_t - r_t)^2}{4\sigma^2}\,dt\Big].
$$

3.2 Dual problem

We now develop the dual problem. The general approach is to use the capital constraint to bound the result of the primal problem from above which gives a related minimization problem. Karatzas and Žitković prove (Proposition 2.13 in [27]) that the capital constraint (3.1) can equivalently be formulated as follows, using the set $\mathcal{D}$ introduced above. The set $\mathcal{A}(x,\mu)$ of admissible consumption rates consists of all processes c such that

$$
\big(C_t := \int_0^t c_u\,\mu(du)\big)_{0\le t\le T}
$$

is a non-negative, non-decreasing, right-continuous and adapted process satisfying

$$E\left[\int_0^T Y_t^R c_t\,\mu(dt)\right] \le x + \langle R, \mathcal{E}_T\rangle \qquad \text{for all } R\in\mathcal{D}. \tag{3.10}$$

With the help of this characterization we can motivate our dual problem by the next calculation. Due to (3.10) we have for all $y>0$ and all $R\in\mathcal{D}$

$$\begin{aligned}\mathcal{U}_Q(c) \le& E_Q\Big[\int_0^T U(t,c_t)\,\mu(dt)\Big] + y\Big(x+\langle R,\mathcal{E}_T\rangle - E\Big[\int_0^T c_t Y_t^R\,\mu(dt)\Big]\Big)\\ \le& E\Big[Z^Q\int_0^T U(t,c_t) - y\,c_t\frac{Y_t^R}{Z_t^Q}\,\mu(dt)\Big] + y\Big(x+\langle R,\mathcal{E}_T\rangle\Big)\\ \le& E\Big[Z^Q\int_0^T V\Big(t,y\frac{Y_t^R}{Z_t^Q}\Big)\,\mu(dt)\Big] + y\Big(x+\langle R,\mathcal{E}_T\rangle\Big).\end{aligned}$$

Hence, we arrive at the dual problem

$$v(y) = \inf_{Z\in\mathcal{Z}}\left(\inf_{R\in\mathcal{D}}\Big(E\Big[Z\int_0^T V\Big(t,y\frac{Y_t^R}{Z_t}\Big)\mu(dt)\Big] + y\langle R,\mathcal{E}_T\rangle\Big)+\gamma(Z)\right). \tag{3.11}$$

Remark 3.9
When considering the inner problem we could also define

$$\tilde v_Q(y) = \inf_{R\in\mathcal{D}_Q} E_Q\left[\int_0^T V\Big(t,yY_t^{R,Q}\Big)\mu(dt)\right] + y\,\langle R,\mathcal{E}_T\rangle$$

where $\mathcal{D}_Q$ is the weak*-closure of the set of supermartingale measures equivalent to Q in $(L^\infty(P))^*$ and $Y^{R,Q}$ is again defined as a version of the density process of the regular part of R but here with respect to Q. Obviously $\mathcal{D}_Q$ can be empty if the market given by the measure Q admits arbitrage. In this case we set the infimum to infinity. With the help of the functions $(\tilde v_Q)_{Q\in\mathcal{Q}}$ we can define an alternative dual function $\tilde v$:

$$\tilde v(y) = \inf_{Q\in\mathcal{Q}}\left(\tilde v_Q(y)+\gamma(Q)\right).$$

Schied defines his dual problem in [38] in this way. In contrast to our setting he is able to show that there is a one-to-one correspondence between the two kinds of dual sets [38, Lemma 4.2]. In our case no similar result holds, i.e. there is no way to decide whether R belongs to $\mathcal{D}$ based on the knowledge of $\mathcal{D}_Q$. This is due to the random endowment (Remark 3.7 implies that the result holds in the case without random endowment). Nevertheless, we can still show that $v(y)=\tilde v(y)$ for all $y>0$. If we define the dual problem via the function $\tilde v$ there are cases where the infima are not attained in contrast to Assertion 5 of Theorem 3.5. This will be illustrated by the example below.

Example 3.10
In this example we will have a unique solution $(\widehat R,\widehat Q)$ to the dual problem for which $\widehat R\notin\mathcal{D}_{\widehat Q}$. We consider all convex combinations of two measures Q^* and P, i.e.

$$\mathcal{Q} = \{Q^\alpha = \alpha P + (1-\alpha)Q^* | 0\le\alpha\le 1\}.$$

The penalty function is given by

$$\gamma(Q) = \begin{cases} \alpha n, & \text{for } Q = Q^\alpha, 0 \le \alpha \le 1 \\ \infty, & \text{otherwise,} \end{cases}$$

where n is a constant we will specify later. Under P, $S_0 = 10$ and S_T takes the values 5 and 15 each with probability 0.5. Under Q^*, $S_0 = 10$ and $S_T = 5$ Q^*-a.s. In order to achieve a continuous time model we set $S_t = S_0$ for all $t < T$. Obviously $Q^* \ll P$.

Assume that

$$\mathcal{E}_T = mI_{\{S_T=5\}}, \qquad \mu = \delta_{\{T\}} \qquad \text{and} \qquad U(t,x) = 2\sqrt{x}$$

(again m will be given later). Hence, $V(t,y) = V(y)$ equals $1/y$. Observe that $\mathcal{D}$ contains all measures $R \ll P$ that satisfy

$$\beta := R[S_T = 5] \ge 0.5$$

and $\mathcal{D}_{Q^*}$ contains only Q^*. Then

$$\begin{aligned} v(y) &= \inf_{\alpha\in[0,1]} \inf_{R\in\mathcal{D}} \Big(E_{Q^\alpha}\Big[\int_0^T V(t, \frac{yY_t^R}{Z_t^\alpha})\, \mu(dt) \Big] + \gamma(Q) + y\langle R, \mathcal{E}_T\rangle \Big) \\ &= \inf_{\alpha\in[0,1]} \inf_{\beta\in[1/2,1]} \Big((1-\frac{\alpha}{2}) V\Big(\frac{y\beta}{1-\alpha/2}\Big) + \frac{\alpha}{2} V\Big(\frac{y(1-\beta)}{\alpha/2}\Big) + \alpha n + y\beta m \Big). \end{aligned}$$

Since we want to compute $v(1)$ we need to consider

$$g(\alpha,\beta) = \frac{(1-\alpha/2)^2}{\beta} + \frac{(\alpha/2)^2}{1-\beta} + \alpha n + \beta m.$$

The partial derivatives of g are

$$g_\alpha(\alpha,\beta) = n - \frac{-(1-\alpha/2)}{\beta} + \frac{\alpha/2}{1-\beta}$$

and

$$g_\beta(\alpha,\beta) = m + -\frac{(1-\alpha/2)^2}{\beta^2} + \frac{(\alpha/2)^2}{(1-\beta)^2}.$$

If we choose $m = 4$ and $n = 2$ both derivatives are positive and therefore the infimum is reached for $\alpha = 0$ and $\beta = 1/2$.

Hence, the solution to the dual problem equals $(Q^0, P) = (Q^*, P)$. Obviously $P \notin \mathcal{D}_{Q^*}$.

In the rest of this section we will prove that the solution to our dual problem exists, and that v equals $\tilde{v}$. First we observe that for $Q_0 \in \mathcal{Q}_e$ of Assumption 3.4 Theorem 3.10 in [27] guarantees that $v_{Q_0}(y) < \infty$ for all $y > 0$ and consequently also $v(y) < \infty$ for all $y > 0$. We will repeatedly make use of a version of Komlós principle of convergence, compare e.g. [7, Lemma A1.1]. To control the behavior of the penalty function we will need the following lemma, which is taken from Schied [38, Lemma 4.1].

Lemma 3.11
For $d \geq 0$ denote the subsets of $\mathcal{Z}$ corresponding to

$$\mathcal{Q}(d) := \{Q \in \mathcal{Q} \,|\, \gamma(Q) \leq d\} \qquad \text{and} \qquad \mathcal{Q}_e(d) := \{Q \in \mathcal{Q}(d) \,|\, Q \sim P\}$$

by $\mathcal{Z}(d)$ and $\mathcal{Z}_e(d)$. Then for every $d > 0$, the level set $\mathcal{Z}(d)$ is weakly compact, and $\mathcal{Z}_e(d)$ is nonempty. Moreover, $Z \mapsto \gamma(Z)$ is lower semicontinuous with respect to P-a.s. convergence on $\mathcal{Z}(d)$.

We will need the following technical result.

Lemma 3.12
For each constant $d > 0$ the set,

$$\Big\{ZV^-\Big(\,\cdot\,, \frac{y\,Y^R_\cdot}{Z_\cdot}\Big) \,|\, Z \in \mathcal{Z}(d), R \in \mathcal{D}\Big\},$$

is uniformly integrable with respect to $P \otimes \mu$.

Proof. Proposition 3.5 in [27] guarantees the existence of a utility function $\underline{U}$ such that

$$\underline{U}(x) \leq U(t, x)$$

for all $x > 0$ and all $t \in [0, T]$. Furthermore, the convex conjugate $\underline{V}$ to $\underline{U}$ satisfies

$$\underline{V}(\cdot) \leq V(t, \cdot)$$

for all $0 \leq t \leq T$. Now we give a slight modification of the arguments of the proof of Lemma 3.6 in [40] to get the claim for $\underline{V}$ and hence, also for V.

Since $\mathcal{Z}(d)$ is uniformly integrable (according to Lemma 3.11 and the Dunford-Pettis theorem) the claim follows immediately if $\underline{V}$ is bounded from below. Assume $\underline{V}$ is not bounded from below. Let φ denote the inverse function of $-\underline{V}$ and $y_0 = \varphi(0)$. Then it follows that

$$\begin{aligned} E\Big[\int_0^T Z_t \varphi\Big(V^-\Big(y\frac{Y^R_t}{Z_t}\Big)\Big)\,\mu(dt)\Big] \leq & E\Big[\int_0^T Z_t \varphi\Big(-\underline{V}\Big(y\frac{Y^R_t}{Z_t}\Big)\Big)\,\mu(dt)\Big] + y_0 \\ \leq & E\left[\int_0^T y\,Y^R_t\,\mu(dt)\right] + y_0 \\ \leq & y + y_0 = M. \end{aligned}$$

In [28] it was proved that $\varphi(h)/h \to \infty$ as $h \to \infty$. Hence for each a there exists $d(a)$ such that $\varphi(h) \geq ah$ for all $h \geq d(a)$. Let $\varepsilon > 0$ and take $d = d(2M/\varepsilon)$, $\eta = \varepsilon/2d$. If

$$A \in \mathcal{F} \otimes \mathcal{B}([0, T]) \quad \text{with} \quad E\Big[\int_0^T Z_t I_A\,\mu(dt)\Big] < \eta$$

then we obtain the following inequality.

$$
\begin{aligned}
&E\Big[\int_0^T Z_t V^-\Big(t,\frac{y\,Y_t^R}{Z_t}\Big) I_A\,\mu(dt)\Big]\\
&\le E\Big[\int_0^T Z_t \underline{V}^-\Big(\frac{y\,Y_t^R}{Z_t}\Big)\Big(I_{A\cap\{\underline{V}^-(y\,Y_t^R/Z_t)\ge d\}}+I_{A\cap\{\underline{V}^-(y\,Y_t^R/Z_t)<d\}}\Big)\,\mu(dt)\Big]\\
&\le E\Big[\int_0^T Z_t\varphi\Big(\underline{V}^-\Big(\frac{y\,Y_t^R}{Z_t}\Big)\Big) I_{A\cap\{\underline{V}^-(y\,Y_t^R/Z_t)\ge d\}}\,\mu(dt)\Big]\,\frac{\varepsilon}{2M}+dE\Big[\int_0^T Z_t I_A\,\mu(dt)\Big]\\
&<\varepsilon.
\end{aligned}
$$

Because of the uniform integrability of $\mathcal{Z}(d)$ there exists $\delta>0$ such that

$$
E\Big[\int_0^T Z_t I_A\,\mu(dt)\Big]<\eta
$$

as soon as $(P\otimes\mu)[A]<\delta$. This finishes the proof. $\square$

The next lemma shows assertion 5 of the main theorem. In this and some later proofs we will need the fact that the map $(x,y)\mapsto xV(t,y/x)$ is convex for all $0\le t\le T$, and that

$$
\alpha\,x_0V(t,y_0/x_0)+(1-\alpha)x_1V(t,y_1/x_1)>(\alpha x_0+(1-\alpha)x_1)V\Big(t,\frac{\alpha y_0+(1-\alpha)y_1}{\alpha x_0+(1-\alpha)x_1}\Big)
$$

for $y_1/x_1\neq y_0/x_0$ and $\alpha\in(0,1)$; see e.g. [38, equation (4.4)].

Lemma 3.13
There exist $\widehat{Z}\in\mathcal{Z}$ and $\widehat{R}\in\mathcal{D}$ such that

$$
v(y)=E\Big[\widehat{Z}\int_0^T V\Big(t,y\frac{Y_t^{\widehat{R}}}{\widehat{Z}_t}\Big)\,\mu(dt)\Big]+y\langle\widehat{R},\mathcal{E}_T\rangle+\gamma(\widehat{Z}).
$$

Proof. Observe that Assumption 3.4 guarantees that $v(y)<\infty$ (compare Karatzas and Žitković [27, Theorem 3.10]). Let $(Z^n,R^n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{Z}\times\mathcal{D}$ such that

$$
E\Big[Z^n\int_0^T V\Big(t,y\frac{Y_t^{R^n}}{Z_t^n}\Big)\,\mu(dt)\Big]+y\langle R^n,\mathcal{E}_T\rangle+\gamma(Z^n)\to v(y)\text{ as }n\to\infty.
$$

Let $dQ^n=Z^ndP$, then Jensen's inequality and Fubini's theorem imply that

$$
\begin{aligned}
E\Big[Z^n\int_0^T V\Big(t,y\frac{Y_t^{R^n}}{Z_t^n}\Big)\,\mu(dt)\Big]&\ge E\Big[Z^n\int_0^T \underline{V}\Big(y\frac{Y_t^{R^n}}{Z_t^n}\Big)\,\mu(dt)\Big]\\
&\ge\int_0^T\underline{V}\Big(E_{Q^n}\Big[\frac{y\,Y_t^{R^n}}{Z_t^n}\Big]\Big)\,\mu(dt)\\
&=\int_0^T\underline{V}\Big(E\Big[y\,Y_t^{R^n};Z_t^n>0\Big]\Big)\,\mu(dt)\\
&\ge\underline{V}(y).
\end{aligned}
\tag{3.12}
$$

Since $\mathcal{E}_T \geq 0$ it follows that

$$d := 1 + \limsup_{n\to\infty} \gamma(Z^n) < \infty$$

and we may assume that $Z^n \in \mathcal{Z}(d)$ for all n.

Now we want to construct a sequence converging to the optimal $(\widehat{Z}, \widehat{R})$. For this we apply repeatedly a Komlós-type argument. Observe that in our case convex combinations of a converging sequence are still convergent. First we obtain a sequence $(\tilde{Z}^n, \tilde{R}^n)_n$ of convex combinations of $(Z^n, R^n)_n$, i.e.

$$(\tilde{Z}^n, \tilde{R}^n) \in \operatorname{conv}\{(Z^n, R^n), (Z^{n+1}, R^{n+1}), \ldots\},$$

where $\tilde{Z}^n$ converges P - a.s. to some $\widehat{Z} \in \mathcal{Z}(d)$ (compare Lemma 3.11). To get also convergence for the processes $(Z_t)_{0\leq t\leq T}$ and $(Y_t^{R^n})_{0\leq t\leq T}$ we will argue analogously to the proof of Proposition A.2 in [27]. Due to [14, Lemma 5.2] we can choose a sequence of convex combinations $((\widehat{Z}_t^n)_{0\leq t\leq T}, (Y_t^{\widehat{R}^n})_{0\leq t\leq T})$ of elements in $((\tilde{Z}_t^m)_{0\leq t\leq T}, (Y_t^{\tilde{R}^m})_{0\leq t\leq T})_{m\geq n}$ that Fatou-converges to some RCLL supermartingale $(\widehat{Z}_t)_{0\leq t\leq T}$. Due to Lemma 3.11 and the fact that μ is diffuse on $[0, T)$ we can find a subsequence that converges also $P \otimes \mu$-a.s. to $(\widehat{Z}_t)_{0\leq t\leq T}$. We know that $\tilde{Z} = \widehat{Z}_T \in \mathcal{Z}(d)$ hence $(\widehat{Z}_t)_{0\leq t\leq T}$ is a density process. To get similarly $Y^{\widehat{R}}$ we first extract a subsequence of $(\widehat{R}^n, \widehat{Z}^n)_n$ also denoted by $(\tilde{R}^n, \tilde{Z}^n)_n$ such that $\langle \widehat{R}^n, \mathcal{E}_T\rangle$ converges in $\mathbb{R}$. Then we consider the corresponding series $(Y^{\widehat{R}^n})_{n\in\mathbb{N}}$. We have a sequence $(\tilde{Y}^{\widehat{R}^n})_{n\in\mathbb{N}}$ of convex combinations of $(Y^{\widehat{R}^n})$ converging $P \otimes \mu$ - a.s. to some $Y^{\widehat{R}}$ where $\widehat{R}$ is a weak* cluster point of $\widehat{R}^n \in \mathcal{D}$.

For $Z \in \mathcal{Z}(d), R \in \mathcal{D}$ the function

$$(Z, Y^R) \mapsto E\Big[\, Z \int_0^T V(t, y\, Y_t^R / Z_t)\, \mu(dt) \Big]$$

is lower semicontinuous with respect to P-a.s. convergence. This can be proved the same way as Lemma 3.7 in [40] using our Lemma 3.12. Combining this lower semicontinuity with Lemma 3.11 and the fact that $(x, y) \mapsto xV(t, y/x)$ is a convex function, results in

$$\begin{aligned} E\Big[\, \widehat{Z} \int_0^T V\Big(t, y\frac{Y_t^{\widehat{R}}}{\widehat{Z}_t}\Big)\, \mu(dt) \Big] &+ y\langle \widehat{R}, \mathcal{E}_T\rangle + \gamma(\widehat{Z}) \\ &\leq \liminf_{n\to\infty} \Big(E\Big[\, Z^n \int_0^T V\Big(t, y\frac{Y_t^{R^n}}{Z_t^n}\Big)\, \mu(dt) \Big] + y\langle R^n, \mathcal{E}_T\rangle + \gamma(Z^n) \Big) \\ &= v(y). \end{aligned}$$

This proves the optimality of $(\widehat{Z}, \widehat{R})$. ☐

We will now show that we may replace the set $\mathcal{Q}$ in the dual problem by $\mathcal{Q}_e$ as well as $\mathcal{Q}_e^f$, and that the dual functions are equal. (Where $\mathcal{Q}_e^f$ denotes the set of measures $Q \in \mathcal{Q}_e$ where $u_Q(x)$ is finite for some $x > 0$.)

Lemma 3.14
The dual value function of the robust problem satisfies

$$\tilde{v}(y) = v(y) = \inf_{Q \in \mathcal{Q}_e^f} \left(v_Q(y) + \gamma(Q) \right) = \inf_{Q \in \mathcal{Q}_e} \left(v_Q(y) + \gamma(Q) \right).$$

Proof. In order to get

$$v(y) = \inf_{Q \in \mathcal{Q}_e} \left(v_Q(y) + \gamma(Q) \right)$$

we adapt the proof of Lemma 4.4 in [38] as follows. Let $(Z^1, R^1) \in \mathcal{Z} \times \mathcal{D}$ be such that

$$v(y) = E\Big[Z^1 \int_0^T V(t, y\, Y_t^{R^1}/Z_t^1)\, \mu(dt) \Big] + \gamma(Z^1) + \langle R^1, \mathcal{E}_T \rangle < \infty.$$

Due to Assumption 3.4 and the assumptions on U we can choose $Z^0 \in \mathcal{Z}_e$ and $R^0 \in \mathcal{D}$ such that also

$$E[\, Z^0 \int_0^T V(t, y\, Y_t^{R^0}/Z_t^0)\, \mu(dt)\,] + \gamma(Z^0) < \infty.$$

Let $Z^\alpha := \alpha Z^1 + (1-\alpha) Z^0 \in \mathcal{Z}_e$ and $R^\alpha := \alpha R^1 + (1-\alpha) R^0$ for $0 \le \alpha < 1$. Since

$$\alpha \mapsto E\Big[Z^\alpha \int_0^T V(t, y\, Y_t^{R^\alpha}/Z_t^\alpha)\, \mu(dt) \Big]$$

is convex and takes only finite values it is upper semicontinuous. Moreover, with the same argument we can conclude that $\alpha \mapsto \gamma(Z^\alpha)$ is upper semicontinuous. This yields together with Lemma 3.11 that $\alpha \mapsto \gamma(Z^\alpha)$ is continuous[2]. The functional $\alpha \mapsto \langle R^\alpha, \mathcal{E}_T \rangle$ is linear and bounded and hence continuous. Consequently, the function

$$\begin{aligned} \alpha \mapsto \inf_{R \in \mathcal{D}} &\Big(E\Big[Z^\alpha \int_0^T V\Big(t, \frac{y\, Y_t^R}{Z_t^\alpha}\Big) \mu(dt) \Big] + y \langle R, \mathcal{E}_T \rangle \Big) + \gamma(Z^\alpha) \\ &= v_{Z^\alpha}(y) + \gamma(Z^\alpha), \end{aligned}$$

is also upper semicontinuous on $[0,1]$, therefore we get

$$\begin{aligned} v(y) &= E\Big[Z^1 \int_0^T V\Big(t, \frac{y\, Y_t^{R^1}}{Z_t^1}\Big) \mu(dt) \Big] + y \langle R^1, \mathcal{E}_T \rangle + \gamma(Z^1) \\ &\ge \limsup_{\alpha \nearrow 1} (v_{Z^\alpha}(y) + \gamma(Z^\alpha)). \end{aligned}$$

This yields

$$v(y) = \inf_{Q \in \mathcal{Q}_e} \left(v_Q(y) + \gamma(Q) \right).$$

Furthermore, observe that

$$v_Q(y) = \infty \quad \text{for} \quad Q \in \mathcal{Q}_e \backslash \mathcal{Q}_e^f.$$

[2] For a more detailed argument compare [38, Remark 4.5].

This follows again from [27, Lemma A.3] for $Q \sim P$ since then v_Q and u_Q satisfy the duality relations. We have $v(y) \leq \tilde{v}(y)$ as $\mathcal{D}_Q \subset \mathcal{D}$ for $Q \ll P$. Since $Z^\alpha \in \mathcal{Z}_e$ for $\alpha \in (0,1)$ we also get

$$v(y) \geq \limsup_{\alpha \nearrow 1}(v_{Z^\alpha}(y) + \gamma(Z^\alpha)) \geq \inf_{Z \in \mathcal{Z}}(v_Z(y) + \gamma(Z)) = \tilde{v}(y).$$

This proves the first identity. □

3.3 Proofs for the primal problem

In this section we will prove the missing assertions. First we make some simple observations. Due to (3.10) we know that for all $\alpha \in [0,1]$ we have

$$\alpha\mathcal{A}(x_1,\mu) + (1-\alpha)\mathcal{A}(x_2,\mu) \subset \mathcal{A}(\alpha x_1 + (1-\alpha)x_2, \mu).$$

Furthermore, if Assumption 3.4 is satisfied it is easy to show that, under the convention (3.5), $c \mapsto \mathcal{U}_Q(c)$ is a concave functional on $\mathcal{A}(x,\mu)$ for each $Q \in \mathcal{Q}$ and all $x > 0$. These facts yield the concavity of the value functions u_Q and u and therefore under Assumption 3.4 the finiteness of the value function $u(x)$ for all $x > 0$. The concavity of u_Q implies in turn that $u_Q \equiv +\infty$ as soon as

$$E_Q\Big[\int_0^T U^+(t,c_t)\,\mu(dt)\Big] = +\infty$$

for some $c \in \bigcup_{x>0}\mathcal{A}(x,\mu)$. Indeed, if

$$E_Q\Big[\int_0^T U^+(t,c_t)\,\mu(dt)\Big] = +\infty \qquad \text{for } c \in \mathcal{A}(x,\mu)$$

then it follows that $c + 1 \in \mathcal{A}(x+1,\mu)$ and

$$E_Q\Big[\int_0^T U(t,c_t)\,\mu(dt)\Big] = +\infty.$$

Thus, $u_Q(x+1) = +\infty$. As u_Q is concave this implies $u_Q \equiv +\infty$.

Lemma 3.15
We have the following minimax identity.

$$\begin{aligned} u(x) &= \sup_{c\in\mathcal{A}(x,\mu)} \inf_{Q\in\mathcal{Q}}\big(\mathcal{U}_Q(c) + \gamma(Q)\big) = \inf_{Q\in\mathcal{Q}}\big(u_Q(x) + \gamma(Q)\big) \\ &= \sup_{c\in\mathcal{A}(x,\mu)} \inf_{Q\in\mathcal{Q}_e}\big(\mathcal{U}_Q(c) + \gamma(Q)\big) = \inf_{Q\in\mathcal{Q}_e}\big(u_Q(x) + \gamma(Q)\big). \end{aligned}$$

Hence, Assertion 2 of Theorem 3.5 is valid.

Proof. The proof of this lemma works along the lines of the proof of Lemma 4.6 in [38]. Let $\varepsilon \in (0, 1)$. Proposition 3.5 in [27] gives the existence of a utility function $\underline{U}$ such that

$$\underline{U}(x) \le U(t, x)$$

for all $x > 0$ and all $t \in [0, T]$. With the help of this utility function we define $d = 1 + u(x+1) - \underline{U}(\varepsilon) \wedge 0$. Then we have

$$\begin{aligned} u(x+1) \ge u(x+\varepsilon) &\ge \sup_{c \in \mathcal{A}(x,\mu)} \inf_{Q \in \mathcal{Q}} \big(\mathcal{U}_Q(c+\varepsilon) + \gamma(Q) \big) \\ &= \sup_{c \in \mathcal{A}(x,\mu)} \inf_{Z \in \mathcal{Z}(d)} \big(\mathcal{U}_Q(c+\varepsilon) + \gamma(Z) \big). \end{aligned}$$

Now $U(t, . + \varepsilon)$ is bounded from below and thus

$$\begin{aligned} Z \mapsto E &\left[Z \int_0^T U(t, c_t + \varepsilon)\, \mu(dt) \right] \\ &= \sup_n E \left[Z \left(\int_0^T U(t, c_t + \varepsilon)\, \mu(dt) \wedge n \right) \right] \end{aligned}$$

is a weakly lower semicontinuous affine functional. Lemma 3.11 states that $Z \mapsto \gamma(Z)$ is weakly lower semicontinuous and $\mathcal{Z}(d)$ is a weakly compact and convex set. Furthermore, for each $Z \in \mathcal{Z}(d)$ $c \mapsto \mathcal{U}_Z(c)$ is a concave functional defined on the convex set $\mathcal{A}(x, \mu)$. Therefore, we may use the lopsided minimax theorem [1, p.295] to obtain

$$\sup_{c \in \mathcal{A}(x,\mu)} \min_{Z \in \mathcal{Z}(d)} \big(\mathcal{U}_Z(c+\varepsilon) + \gamma(Z) \big) = \min_{Z \in \mathcal{Z}(d)} \sup_{c \in \mathcal{A}(x,\mu)} \big(\mathcal{U}_Z(c+\varepsilon) + \gamma(Z) \big).$$

We know that these expressions are bounded by $u(x+\varepsilon) < d + \underline{U}(\varepsilon) \wedge 0$. Thus, it does not matter whether we take the infimum over $\mathcal{Z}$ or over $\mathcal{Z}(d)$. We obtain

$$u(x+\varepsilon) \ge \inf_{Z \in \mathcal{Z}} \sup_{c \in \mathcal{A}(x,\mu)} \big(\mathcal{U}_Z(c+\varepsilon) + \gamma(Z) \big) \tag{3.13}$$

$$\ge \inf_{Z \in \mathcal{Z}} \sup_{c \in \mathcal{A}(x,\mu)} \big(\mathcal{U}_Z(c) + \gamma(Z) \big) \tag{3.14}$$

$$\ge \sup_{c \in \mathcal{A}(x,\mu)} \inf_{Z \in \mathcal{Z}} \big(\mathcal{U}_Z(c) + \gamma(Z) \big) = u(x). \tag{3.15}$$

As $\varepsilon \to 0$ the assertion follows since u is continuous. □

Observe now that Theorem 3.10 in [27] states that

$$u_Q(x) = \inf_{y>0} (v_Q(y) + xy) \quad \text{for all} \quad Q \in \mathcal{Q}_e^f.$$

Pooling our lemmas and using this result, we get

$$\begin{aligned} u(x) &= \inf_{Q \in \mathcal{Q}_e} (u_Q(x) + \gamma(Q)) = \inf_{Q \in \mathcal{Q}_e^f} (u_Q(x) + \gamma(Q)) \\ &= \inf_{Q \in \mathcal{Q}_e^f} \inf_{y>0} (v_Q(y) + \gamma(Q) + xy) = \inf_{y>0} (v(y) + xy) \end{aligned}$$

which is assertion 3 of the theorem. Finiteness of v and general duality principles yield then

$$\lim_{x\to\infty} u(x)/x \to 0. \tag{3.16}$$

Lemma 3.16
For any $x > 0$ there exists $\widehat{c} \in \mathcal{A}(x,\mu)$ such that

$$\inf_{Q\in\mathcal{Q}}\left(E_Q\Big[\int_0^T U(t,\widehat{c}_t)\,\mu(dt) \Big] + \gamma(Q) \right) = u(x).$$

Proof. Let $(c^n)_{n\in\mathbb{N}}$ be a maximizing sequence for a given $x > 0$. Using again a Komlós-type argument we get a sequence $(\tilde{c}^n)_{n\in\mathbb{N}}$ with $\tilde{c}^n \in \operatorname{conv}(c^n, c^{n+1}, \ldots)$ converging $P\otimes\mu$-a.s. to some $\widehat{c} \in \mathcal{A}(x,\mu)$ since $\mathcal{A}(x,\mu)$ is closed under convergence in probability.

The following adaption of an argument in the proof of [29, Lemma 1] shows that the positive parts of $U(t,\tilde{c}^n_t)$ are uniformly integrable with respect to $Q\otimes\mu$ for all $Q \in Q^f_e$.

Assume $(U^+(t,\tilde{c}^n_t))_{n\in\mathbb{N}}$ is not uniformly integrable. Then there is a constant α, a subsequence which is also denoted $(\tilde{c}_n)_{n\in\mathbb{N}}$ and a disjoint partition $(A_n)_{n\in\mathbb{N}}$ of $(\Omega\times[0,T], \mathcal{F}\otimes\mathcal{B}([0,T]))$ such that

$$E\Big[\int_0^T U^+(t,\tilde{c}^n_t) I_{A_n}\,\mu(dt) \Big] \geq \alpha \text{ for } n \geq 1.$$

For $0 \le t \le T$ let

$$x^0_t = \inf\{x > 0 \,|\, U(t,x) \geq 0\}$$

and define a process $s^n = (s^n_t)_{0\le t\le T}$ by

$$s^n_t = x^0_t + \sum_{k=1}^n \tilde{c}^k_t I_{A_k}.$$

Then for any $Y \in \mathcal{Y}$ we have

$$E\Big[\int_0^T Y_t s^n_t\,\mu(dt) \Big] \leq \int_0^T x^0_t\,\mu(dt) + nx.$$

Hence $s^n \in \mathcal{A}\big(\int_0^T x^0_t\,\mu(dt) + nx, \mu\big)$. Furthermore

$$E\Big[\int_0^T U(t,s^n_t)\,\mu(dt) \Big] \geq \sum_{k=1}^n E\Big[\int_0^T U^+(t,\tilde{c}^k_t) I_{A_k}\,\mu(dt) \Big] \geq \alpha n.$$

Which yields

$$\begin{aligned}
\limsup_{x\to\infty} \frac{u(x)}{x} &\geq \limsup_{n\to\infty} \frac{E\big[\int_0^T U(t,s_t)\,\mu(dt) \big]}{\int_0^T x^0_t \mu(dt) + nx} \\
&\geq \limsup_{n\to\infty} \frac{\alpha n}{\int_0^T x^0_t \mu(dt) + nx} \\
&= \alpha > 0.
\end{aligned}$$

Contradicting (3.16).

Using the uniform integrability on the positive parts and a version of Fatou's Lemma on the negative parts of $U(t, \tilde{c}_t^n)$ we can deduce that

$$c \mapsto E_Q\left[\int_0^T U(t, c_t)\,\mu(dt)\right] + \gamma(Q)$$

is upper semicontinuous with respect to almost sure convergence. From the concavity of

$$c \mapsto \inf_{Q \in \mathcal{Q}_e^f} \Big(E_Q\Big[\int_0^T U(t, c_t)\,\mu(dt)\Big] + \gamma(Q) \Big)$$

it follows that $(\tilde{c}^n)$ is still a maximizing sequence. Then the upper semicontinuity yields that

$$\inf_{Q \in \mathcal{Q}_e} \Big(E_Q\Big[\int_0^T U(t, \widehat{c}_t)\,\mu(dt)\Big] + \gamma(Q) \Big) \geq u(x).$$

Actually, we also have

$$\inf_{Q \in \mathcal{Q}} \Big(E_Q\Big[\int_0^T U(t, \widehat{c}_t)\,\mu(dt)\Big] + \gamma(Q) \Big) \geq u(x):$$

first we have that

$$\{\, Q \in \mathcal{Q} \,|\, E_Q[\int_0^T U(t, \widehat{c}_t)\,\mu(dt)\,] = -\infty \,\} = \emptyset$$

since otherwise this set would have a nonempty intersection with $\mathcal{Q}_e$. Hence, for $Q \in \mathcal{Q} \backslash \mathcal{Q}_e$, $Q_0 \in \mathcal{Q}_e^f$ and

$$Q_\alpha := \alpha Q + (1 - \alpha) Q_0 \in \mathcal{Q}_e,$$

we get

$$\lim_{\alpha \nearrow 1} E_{Q_\alpha}\Big[\int_0^T U(t, \widehat{c}_t)\,\mu(dt)\Big] = E_Q\Big[\int_0^T U(t, \widehat{c}_t)\,\mu(dt)\Big].$$

Due to the convexity and lower semicontinuity of γ, we also have $\gamma(Q_\alpha) \to \gamma(Q)$. $\square$

Proof of Theorem 3.5. Assertion 2 and 3, the finiteness of u and v, and the concavity of u were already proved. The convexity of v is an immediate consequence of the convexity of

$$(Z, y\,Y^R) \mapsto \mathcal{V}_Z(Y^R/Z) + \gamma(Z).$$

Assertion 3 of Theorem 3.5 is a consequence of the lower semicontinuity of v (the lower semicontinuity can be shown as in the proof of Theorem 2.3 in [38] with an additional $\langle R, \mathcal{E}_T \rangle$-term).

Equation (3.8) follows as in [27, Lemma A.7]. If $\mathcal{E}_T \equiv 0$ this implies $v'(\infty-) = 0$ which yields by general duality results that $u'(0) = \infty$ which is equation (3.9). Assertion 5 corresponds to Lemma 3.13. Furthermore, the existence of an optimal c (Item 6) is the content of Lemma 3.16.

We will now deal with the existence of a saddlepoint. Let $y > 0$ be such that

$$v(y) + yx = u(x),$$

such a y exists due to the behavior of v'. Take then a solution $(\widehat{Q}, \widehat{R})$ to the dual problem for y and let $\widehat{Z}$ be the density of $\widehat{Q}$ with respect to P. Let Z^1 be in $\mathcal{Z}_e^f$ and define

$$Z^\alpha = \alpha Z^1 + (1-\alpha)\widehat{Z} \quad \text{for} \quad \alpha \in (0,1].$$

Then

$$v_{Z^\alpha}(y) + \gamma(Z^\alpha) \to v(y) \quad \text{as} \quad \alpha \to 0 :$$

Let R^1 be such that

$$\mathcal{V}_{Z^1}(yY^{R^1}/\widehat{Z}^1) + \langle R^1, \mathcal{E}_T\rangle = v_{Z^1}(y).$$

Then define $R^\alpha = \alpha R^1 + (1-\alpha)\widehat{R}$. We get

$$\begin{aligned} v(y) &\leq v_{Z^\alpha}(y) + \gamma(Z^\alpha) \\ &\leq \mathcal{V}_{Z^\alpha}(yY^{R^\alpha}/Z^\alpha) + \langle R^\alpha, \mathcal{E}_T\rangle + \gamma(Z^\alpha) \\ &\leq \alpha\big(v_{Z^1}(y) + \gamma(Z^1)\big) + (1-\alpha)\big(\tilde{v}_{\widehat{Z}}(y) + \gamma(\widehat{Z})\big) \end{aligned}$$

because of convexity. Observe that the right-hand side of the equation goes to $v(y)$ as α goes to 0.

Due to the duality relations we have that

$$v_{Z^\alpha} + xy \geq u_{Z^\alpha} \quad \text{and} \quad u_{Z^\alpha} + \gamma(Z^\alpha) \to u_{\widehat{Z}} + \gamma(\widehat{Z}).$$

Hence,

$$\begin{aligned} u(x) = v(y) + yx &= \lim_{\alpha\to 0}\big(v_{Z^\alpha}(y) + xy + \gamma(Z^\alpha)\big) \\ &\geq \lim_{\alpha\to 0}\big(u_{Z^\alpha}(x) + \gamma(Z^\alpha)\big) = u_{\widehat{Z}}(x) + \gamma(\widehat{Z}) \end{aligned}$$

With the minimax identity we get $u(x) = u_{\widehat{Z}} + \gamma(\widehat{Z})$ and therefore we get for $\hat{c}$ as in Lemma 3.16

$$u(x) = u_{\widehat{Z}}(x) + \gamma(\widehat{Z}) \geq \mathcal{U}_{\widehat{Z},\mu}(\hat{c}) + \gamma(\widehat{Z}) \geq \inf_{Q\in\mathcal{Q}} \mathcal{U}_Q(\hat{c}) + \gamma(\widehat{Z}) = u(x).$$

Now we show that

$$\hat{c} = I(\cdot, yY^{\widehat{R}}) \quad \widehat{Q}\otimes\mu - \text{a.s.}$$

We have

$$0 \leq V\Big(t, \frac{\widehat{y}\,Y_t^{\widehat{R}}}{\widehat{Z}_t}\Big) + \frac{\widehat{y}\,Y_t^{\widehat{R}}}{\widehat{Z}_t}\,\hat{c}_t - U(t, \hat{c}_t)$$

and therefore

$$
\begin{aligned}
0 \leq & E_{\widehat{Q}}\Big[\int_0^T V\Big(t, \frac{y\,Y_t^{\widehat{R}}}{\widehat{Z}_t}\Big)\,\mu(dt) + \int_0^T \frac{y\,Y_t^{\widehat{R}}}{\widehat{Z}_t}\,\hat{c}_t\,\mu(dt) - \int_0^T U(t,\hat{c}_t)\,\mu(dt)\Big] \\
= & v(y) + E\Big[\int_0^T y\,Y_t^{\widehat{R}}\,\hat{c}_t\,\mu(dt)\Big] - y\langle \widehat{R}, \mathcal{E}_T\rangle - u(x) \\
\leq & v(y) + yx - u(x) = 0.
\end{aligned}
$$

Together this implies

$$
0 = V\Big(t, \frac{y\,Y_t^{\widehat{R}}}{\widehat{Z}_t}\Big) + \frac{y\,Y_t^{\widehat{R}}}{\widehat{Z}_t}\hat{c}_t - U(t,\hat{c}_t) \qquad \widehat{Q}\otimes\mu(dt)\text{-a.s.}
$$

which means that $\hat{c}_t = I(t, y\,Y_t^{\widehat{R}}/\widehat{Z}_t)$.

The fact that

$$
u'(\infty-) = 0
$$

follows from (3.16). Therefore, we also have

$$
v'(0+) = -\infty.
$$

To obtain the strict concavity of the function u (and hence the differentiability of v) assume that u is not strictly concave. Since u is increasing and because of the conditions for the derivatives we know there exist $0 < x_0 < x_1$ and $y > 0$ such that

$$
v(y) + yx_0 = u(x_0) \quad \text{ and } \quad v(y) + yx_1 = u(x_1).
$$

Let c^0, c^1 be the corresponding optimal consumption processes and $(\widehat{Q}, \widehat{R})$ the solution to the dual problem. Then we have

$$
c_t^0 = I(t, yY^{\widehat{R}}) = c_t^1 \quad \widehat{Q} - \text{a.s.}
$$

and

$$
E_{\widehat{Q}}\Big[\int_0^T c_t^0 yY_t^{\widehat{R}}\mu(dt)\Big] + y\langle \widehat{R}, \mathcal{E}_T\rangle = x_0 y < x_1 y = E_{\widehat{Q}}\Big[\int_0^T c_t^1 yY_t^{\widehat{R}}\mu(dt)\Big] + y\langle \widehat{R}, \mathcal{E}_T\rangle.
$$

This is a contradiction. □

Proof of Corollary 3.6. This corollary can be proved by copying the arguments of the proof of Proposition 2.4 in [38]. □

Chapter 4

Application to a Markovian switching Model

In the following we want to apply the results of the first part to a specific market model and coherent utility functional given by a specific ambiguity set $\mathcal{Q}$ and a HARA utility function. As a reward for this limitation we can describe the solution to the optimization problem as the solution to the Hamilton-Jacobi-Bellman equations of stochastic optimal control. In our case these are ordinary differential equations and can be solved numerically. Then it is also possible to compute the optimal investment strategy.

As market model we will consider a Markovian switching model. The state of the economy is given by a continuous time Markov chain. We assume that this state influences the coefficients of a simple jump diffusion model and is known to the agent. In the literature there are several papers on Markovian switching models with unobservable state process. In these cases a suitable filter needs to be applied. As far as we know there is no work on filtering in a situation subject to model uncertainty. This poses new questions of interpretation: The filter is applied because some parameter is not observable. This lack of information is also the source of ambiguity. The filtering technique gives a likely parameter while the robust approach also gives some parameter (here, the worst case). Hence, the question arises which part of the uncertainty should be removed via filtering and which part with the help of the robust approach. Or should the robust approach be applied after filtering? Or how else can the situation be resolved? We are not going to answer these questions and therefore restrict ourselves to the case where the state process is observable.

The investor is ambiguous about the rate with which the Markov chain changes its state. We consider HARA utility functions with risk aversion parameter $\alpha \geq 0$.

First we introduce our model and prove that our set $\mathcal{Q}$ satisfies the assumptions of the first part. Since we want to work with the dual problem we need the dual set which we consider in Section 4.2. Then we deal with optimal consumption for an investor with HARA utility where we consider logarithmic utility and HARA utility with risk aversion coefficient $0 < \alpha < 1$ in separate sections. First we study the problem for the logarithmic utility. Here, we develop the HJB equations, give some properties of their solution and compute the optimal strategy. This is done for a *pessimistic* and an *optimistic* investor. We call an investor pessimistic if she optimizes according to MEU, i.e. the agent maximizes

$$\inf_{Q\in\mathcal{Q}} E_Q\Big[\int_0^T \log(c_t)\,\mu(dt)\Big].$$

As before, a possible interpretation is that the investor is aware of different scenarios and tries to be on the safe side by assuming that the worst-case-scenario will occur. This kind of investor might also be described as ambiguity averse. In contrast to the pessimistic investor the optimistic or ambiguity loving investor is going to optimize

$$\sup_{Q\in\mathcal{Q}} E_Q\Big[\int_0^T \log(c_t)\,\mu(dt)\Big].$$

This means the investor assumes the best-case-outcome. As mentioned in the introduction it is also possible to consider a mixture of these two extremes but we restrict ourselves to the two counterparts. Furthermore, we give some numerical results regarding the question which approach is more useful in a "real" setting, i.e. we compute the outcome to the strategies of an optimistic/pessimistic investor and compare these as the "real" parameter of the stock price varies.

In the Section 4.5 we consider HARA utility for $0 < \alpha < 1$. As before, we develop the HJB equation, give some properties and compute the optimal strategy.

4.1 Notations

We work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\le T}, P)$ where Ω is the path space of (W, Y). Here, W is a Brownian motion and Y is a time-continuous Markov chain with state space $\{e_1, \dots, e_n\}$ and generator

$$A = \begin{pmatrix} -\lambda^0 & \lambda^0 p_{1,2} & \dots & & \lambda^0 p_{1,n} \\ & \ddots & & & \\ & & & \ddots & \\ \lambda^0 p_{n,1} & \dots & & \lambda^0 p_{n,n-1} & -\lambda_t \end{pmatrix},$$

where, λ^0 is the fixed jump rate of the Markov chain and $\mathbb{P} = (p_{i,j})_{i,j=1,\dots,n}$ is a stochastic matrix (with $p_{i,i} = 0$ for $i = 1, \dots, n$). Hence, we can equivalently describe Y through a Poisson process N with intensity $\lambda^0 > 0$ giving the jump times and a discrete time Markov chain $\tilde{Y}$ with transition matrix $\mathbb{P}$ that specifies the transition probabilities if a jump occurs. In the following we will frequently use this equivalent formulation. For the proofs and for some results we restrict ourselves to the case $n = 2$ in order to avoid notational inconveniences.

The filtration is assumed to satisfy the usual conditions. Furthermore, M is the compensated Poisson process, i.e. $M_t = N_t - \lambda^0 t$. Observe that under a probability measure Q with density

$$\frac{dQ}{dP} = \mathcal{E}\Big(\int_0^{\cdot} (\lambda_s - \lambda^0)/\lambda^0 \, dM_s\Big)_T$$

the process $(N_t)_{0\le t\le T}$ is a Poisson process with stochastic intensity $(\lambda_t)_{0\le t\le T}$. ($\mathcal{E}$ denotes the stochastic exponential.)

The dynamics of the stock and the bond depend on the external market factor Y:

$$\begin{aligned} dS_t &= S_{t-}\big(\sigma(Y_t)\,dW_t + b(Y_t)\,dt + \delta(Y_{t-})\,dN_t\big)\\ dS^0_t &= S^0_t r(Y_t)\,dt \end{aligned}$$

where we assume $\sigma > 0$. A solution to this SDE exists. See [35, Theorem V.6].

The investor can invest the initial capital $x > 0$ into the stock and bond. The fraction of her wealth invested in the stock at time t will be denoted by π_t. As before the consumption process $C = (C_t)_{0\le t\le T}$ is assumed to be a non-negative, non-decreasing and adapted RCLL process. More specifically we deal only with consumption processes that can be written as

$$C_t = \int_0^t c_s\,\mu(ds)$$

where

$$\mu(ds) = \kappa(\gamma e^{-\rho t}ds + \delta_T(ds)). \tag{4.1}$$

(With $\gamma \ge 0$, $\rho \ge 0$ and κ such that μ is a probability measure on $[0, T]$.)

The wealth process develops according to the following SDE

$$dX^\pi_t = \frac{X^\pi_{t-}\pi_{t-}}{S_{t-}}dS_t + \frac{X^\pi_{t-}(1-\pi_{t-})}{S^0_{t-}}dS^0_t - c_t\,\mu(dt) \text{ and } X_0 = x.$$

As before we require the wealth process to stay non-negative:

$$X^\pi_t \ge 0 \qquad P-\text{a.s.} \tag{4.2}$$

We denote by $\mathcal{A}(x,\mu)$ the set of consumption process densities $(c_t)_{0\le t\le T}$, for which an admissible strategy exists, such that condition (4.2) is fulfilled.

In this setting we solve the optimization problem of a pessimistic investor

$$\max_{c\in\mathcal{A}(x)} \inf_{Q\in\mathcal{Q}} E_Q\Big[\int_0^T U(c_t)\mu(dt)\Big]$$

for U being a HARA utility function and μ defined as in (4.1).

We consider an investor who is not certain how fast the states of the economy switch, i.e. she is ambiguous about the rate of the Poisson process N. (Even though this is not very realistic, we assume that she knows the transition probabilities. However, in a two-state-economy these are not important.) If we let a_1, a_2 be positive constants then the set $\mathcal{Q}$ appearing in the definition of the robust utility functional is given as

$$\mathcal{Q} = \Big\{Q\,\Big|\,\frac{dQ}{dP} = \mathcal{E}\Big(\int_0 \frac{\lambda_s - \lambda^0}{\lambda^0}\,dM_s\Big) \text{ where } \lambda \in \Lambda\Big\}.$$

Where the set Λ parameterizing the possible integrands is defined as

$$\Lambda = \{\lambda \,|\, \lambda \text{ is a predictable process and } \lambda_s \in [a_1, a_2], a_1 < \lambda^0 < a_2\}.$$

For the logarithmic utility we treat also the problem of an optimistic investor, i.e. solve

$$\max_{c \in \mathcal{A}(x)} \sup_{Q \in \mathcal{Q}} E_Q\Big[\int_0^T \log(c_t)\mu(dt)\Big].$$

If not mentioned otherwise we consider the pessimistic investor.

To use the duality result as given in the first chapter we need the following lemma which guarantees that Assumption 3.1 is satisfied.

Lemma 4.1
The coherent risk measure defined through

$$\rho(Y) = \sup_{Q \in \mathcal{Q}} E_Q[-Y]$$

is sensitive in the sense that $\rho(Y)$ is strictly positive for all $Y \in L^\infty_- \setminus \{0\}$ and continuous from below, i.e. for a sequence $(Y_n) \subset L^\infty$ increasing a.s. to some $Y \in L^\infty$, we have $\rho(Y_n) \searrow \rho(Y)$.

Proof. First we show that $\mathcal{Q}$ is convex. Let $\varepsilon \in (0,1)$ and $Q^1, Q^2 \in \mathcal{Q}$. We want to show that

$$Q := \varepsilon Q^1 + (1-\varepsilon)Q^2 \in \mathcal{Q}.$$

We assume that

$$\frac{dQ^i}{dP} = Z^i_T = \mathcal{E}\Big(\int_0^\cdot \frac{\lambda^i_s - \lambda^0}{\lambda^0}\, dM_s\Big)_T$$

for $\lambda^i \in \Lambda$ and $i = 1, 2$. Now we define

$$Z = \varepsilon Z^1 + (1-\varepsilon)Z^2.$$

Then Z is as Z^1, Z^2 a strictly positive uniformly integrable martingale to which we apply now Itô's formula.

$$\begin{aligned} dZ_t =& \varepsilon Z^1_{t-}\frac{\lambda^1_t - \lambda^0}{\lambda^0}\, dM_t + (1-\varepsilon)Z^2_{t-}\frac{\lambda^2_t - \lambda^0}{\lambda^0}\, dM_t \\ =& Z_{t-}\Big(\varepsilon \frac{Z^1_{t-}}{Z_{t-}}\frac{\lambda^1_t - \lambda^0}{\lambda^0} + (1-\varepsilon)\frac{Z^2_{t-}}{Z_{t-}}\frac{\lambda^2_t - \lambda^0}{\lambda^0}\Big)\, dM_t. \end{aligned}$$

Hence, it follows

$$Z_t = \mathcal{E}\Big(\int_0^\cdot \frac{\tilde{\lambda}_s - \lambda^0}{\lambda^0}\, dM_t\Big)_t.$$

Where,

$$\tilde{\lambda}_s = \varepsilon \frac{Z^1_{t-}}{Z_{t-}}\lambda^1_s + (1-\varepsilon)\frac{Z^2_{t-}}{Z_{t-}}\lambda^2_s \in [a_1, a_2].$$

Furthermore, $\tilde{\lambda}$ is predictable and thus $\tilde{\lambda} \in \Lambda$ which implies $Z = dQ/dP$ is such that $Q \in \mathcal{Q}$.

Next, sensitivity follows from the fact that $P \in \mathcal{Q}$. Continuity from below follows from [30, Lemma 2], [15, Corollary 4.35] and the Dunford-Pettis Theorem once we have shown that

$$\mathcal{Z} = \{dQ/dP \mid Q \in \mathcal{Q}\}$$

is uniformly integrable and closed in L^1.

First we show that there is $B \in L^2_+(P)$ such that $dQ/dP \leq B$ for all $Q \in \mathcal{Q}$.

$$\begin{aligned}
\frac{dQ}{dP} &= \exp\Big(-\int_0^T \frac{\lambda_s - \lambda^0}{\lambda^0}\lambda^0 \, ds\Big) \prod_{s \leq T}\Big(1 + \frac{\lambda_s - \lambda^0}{\lambda^0}\Delta N_s\Big) \\
&\leq \exp\big(-\int_0^T (a_1 - \lambda^0)\, ds\big)\Big(\frac{a_2}{\lambda^0}\Big)^{N_T} \\
&=: B
\end{aligned}$$

where

$$\begin{aligned}
E[B^2] &= \exp\big(-\int_0^T 2(a_1 - \lambda^0)\, ds\big) \sum_{n=0}^{\infty} \Big(\frac{a_2}{\lambda^0}\Big)^{2n} P[N_T = n] \\
&= \exp(2(\lambda^0 - a_1)T) e^{-T\lambda^0} e^{a_2^2 T} = e^{T(\lambda^0 + a_2^2 - 2a_1)} < \infty.
\end{aligned}$$

Hence, $\mathcal{Z}$ is uniformly integrable. (In the same way one can show that $B \in L^p(P)$ for $1 \leq p < \infty$.) Furthermore, later we will need a lower bound which we denote by $\tilde{B}$ and that is defined via

$$\tilde{B}_t = \exp\Big(-\int_0^t (a_2 - \lambda^0)\, ds\Big)\Big(\frac{a_1}{\lambda^0}\Big)^{N_t} \text{ for } 0 \leq t \leq T.$$

Obviously,

$$\frac{1}{\tilde{B}_t} \geq \frac{1}{dQ/dP|_{\mathcal{F}_t}} \text{ for all } t \in [0, T].$$

Next we show that $\mathcal{Z}$ is closed in L^1. Let $(Z^n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{Z}$ converging in L^1 to Z. Let

$$Z^n = \frac{dQ^n}{dP} = \mathcal{E}\Big(\int \frac{\lambda^n_s - \lambda^0}{\lambda^0}\, dM_s\Big)_T \quad \text{and} \quad Z = \frac{dQ}{dP}$$

where $\lambda^n \in \Lambda$. We want to show that there exists a predictable process $\lambda^* \in \Lambda$ such that

$$Z = \mathcal{E}\Big(\int \frac{\lambda^*_s - \lambda^0}{\lambda^0}\, dM_s\Big)_T.$$

First we will show that the densities converge also in the semimartingale topology. This topology is induced by the following distance function on the space of semimartingales

$$D[S] = \sum_{n=1}^{\infty} 2^{-n} \sup\Big\{E\Big[\Big|\int_0^n K_u\, dS_u\Big| \wedge 1\Big] \,\Big|\, K \text{ real-valued, predictable, } |K| \leq 1\Big\}.$$

(Compare (7.13) of [8]). Then we can show convergence of the stochastic logarithms and finally we will use Komlós' Theorem to prove the existence of a limit that has the desired form.

We have

$$\|Z^n_T\|_{L^2} \leq \|B\|_{L^2} < c$$

uniformly in n. Hence, Lebesgue's dominated convergence theorem implies that Z^n_T converges to Z_T in L^2 as $n \to \infty$. And this means that $Z^n \to Z$ in $\underline{\underline{H}}^2$. (For a local martingale X the $\underline{\underline{H}}^2$-norm is given by

$$\|X\|_{\underline{\underline{H}}^2} = \|[X,X]^{1/2}_\infty\|_{L^2}.$$

Compare e.g. [35, corollary, p.251].) Furthermore, this convergence yields convergence in the semimartingale topology (compare [35, Theorem IV.14]). We will use this convergence to get also convergence of the stochastic logarithms.[1]

To this end, we first let σ_m be the time of the m-th jump of the Poisson process N. Then $(Z^n)^{\sigma_m}$ is uniformly bounded away from 0, implying that $(1/Z^n)^{\sigma_m}$ and $(1/Z)^{\sigma_m}$ are bounded.

Now we want to show that the stochastic logarithm of Z^n

$$\frac{1}{Z^n_-} \cdot Z^n$$

converges to the stochastic logarithm of Z

$$\frac{1}{Z_-} \cdot Z.$$

(Here, $\cdot$ denotes as usual the stochastic integral.) For this we consider the convergence of the stopped processes first. We have

$$\begin{aligned}\Big|\Big(\frac{1}{Z^n_-}\Big)^{\sigma_m} \cdot Z^n - \Big(\frac{1}{Z_-}\Big)^{\sigma_m} \cdot Z\Big| \leq \\ \Big|\Big(\frac{1}{Z^n_-}\Big)^{\sigma_m} \cdot Z^n - \Big(\frac{1}{Z^n_-}\Big)^{\sigma_m} \cdot Z\Big| + \Big|\Big(\frac{1}{Z^n_-}\Big)^{\sigma_m} \cdot Z - \Big(\frac{1}{Z_-}\Big)^{\sigma_m} \cdot Z\Big|.\end{aligned} \tag{4.3}$$

Here, the first term goes to 0 uniformly on compacts in probability (ucp) since the integrand is bounded and Z^n converges to Z in the semimartingale topology (this is a direct consequence of the definition via (4.3)). The second goes to 0 ucp due to the Theorem of dominated convergence for stochastic integrals, e.g. [35, Theorem IV.32]. As bounding process we can use $1/\tilde{B}_-$ which is in $L(Z^n)$ and $L(Z)$ due

[1] [26, Theorem II.8.3] guarantees the existence of the stochastic logarithm for a real-valued semimartingale Y where Y and Y_- do not vanish. The stochastic logarithm of Y is given as

$$X = \int \frac{1}{Y_-}\, dY.$$

Then X as above is the unique semimartingale that satisfies $Y = \mathcal{E}(X)$ and $X_0 = 0$.

to [35, Theorem IV.28]. Now we let m and hence also σ_m go to infinity, then the left-hand side of (4.3) goes to

$$\Big|\frac{1}{Z^n_-}\cdot Z^n - \frac{1}{Z_-}\cdot Z\Big|$$

(again Theorem of dominated convergence for stochastic integrals). Hence, it follows convergence of the stochastic logarithms.

Let $L = 1/Z_- \cdot Z$. We know that

$$\frac{1}{Z^n_-}\cdot Z^n = \int_0 \frac{\lambda^n_s - \lambda^0}{\lambda^0}\, dM_s.$$

Komlós' Theorem implies that there exists a sequence $(\tilde{\lambda}^n)_{n\in\mathbb{N}}$ of convex combinations of $(\lambda^n)_{n\in\mathbb{N}}$ converging $P\otimes dt$-a.s. to some predictable λ^* with values in $[a_1, a_2]$. Then we have

$$\int_0^t \frac{\lambda^*_s - \lambda^0}{\lambda^0}\, dM_s = \lim_{n\to\infty}\int_0^t \frac{\tilde{\lambda}^n_s - \lambda^0}{\lambda^0}\, dM_s = \lim_{n\to\infty}\int_0^t \frac{\lambda^n_s - \lambda^0}{\lambda^0}\, dM_s = L_t.$$

The first equality follows again from the Theorem of dominated convergence for stochastic integrals. □

4.2 Dual Set

In this chapter $\tilde{U}$ denotes the dual function to U. In order to apply the duality theory we need the corresponding dual set. We use the set

$$\mathcal{D} = \Big\{\mathcal{E}\Big(-\int \frac{b(Y_t) - r(Y_t) + \delta(Y_t)\nu_t}{\sigma(Y_t)}\, dW_t + \int \frac{\nu_t - \lambda^0}{\lambda^0}\, dM_t\Big)_T \,\Big|\, \nu\in\mathcal{N}\Big\}$$

where

$$\mathcal{N} = \Big\{\nu \,\Big|\, \nu_s > 0 \text{ a.e.}, \int_0^t \nu_s^2\, ds < \infty \text{ a.e. and } \nu \text{ predictable}\Big\}.$$

Theorem 4.2
If Assumption 3.2 and Assumption 3.4 are satisfied the dual problem is given as

$$v(y) = \inf_{Q\in\mathcal{Q}}\inf_{D\in\mathcal{D}} E_Q\Big[\int_0^T \tilde{U}\Big(y\frac{D_t}{S^0_t Z^Q_t}\Big)\,\mu(dt)\Big].$$

Hence, we have

$$u(x) = \inf_{y>0}(v(y) + xy).$$

Proof. We will show that $\mathcal{P} \subseteq \mathcal{D} \subseteq \mathcal{Y}(1)$ where $\mathcal{P}$ is the set of equivalent local martingale measures and $\mathcal{Y}(1)$ is the dual set as used e.g. in Remark 3.7. Then the result follows from Remark 3.7 and [38, Remark 2.7] (if we use Lemma 3.14).

To see that $\mathcal{P} \subseteq \mathcal{D}$ let $P^* \in \mathcal{P}$. Then the density process

$$D = (dP^*/dP|_{\mathcal{F}_t})_{0\leq t\leq T}$$

is a strictly positive martingale. Hence, we can take the stochastic logarithm

$$L_t := \int_0^t \frac{1}{D_{s-}} dD_s \qquad \text{for } 0 \leq t \leq T$$

where L is a local martingale. Now we apply a martingale representation result (e.g. [26, III.4.34]) to get

$$L_t = \int \theta_t^\nu \, dW_t + \int \tilde{\nu}_t \, dM_t$$

for $0 \leq t \leq T$ where θ^ν, $\tilde{\nu}$ are predictable processes such that the integrals are well-defined. It follows

$$\frac{dP^*}{dP} = \mathcal{E}\Big(\int \theta_t^\nu \, dW_t + \int \tilde{\nu}_t \, dM_t\Big)_T.$$

In particular $\tilde{\nu} > -1$ and $\int_0^t \tilde{\nu}_s(\omega)^2 \, ds < \infty$ for all $t \in [0,T]$ and P-a.a. $\omega \in \Omega$.

We know that the discounted stock price process $\tilde{S}$ is a local martingale under P^*. Due to Girsanov's Theorem (e.g. [10, Theorem 13.19]) it follows that

$$d\tilde{S}_t = \tilde{S}_t\big(\sigma \, d\tilde{W}_t + \delta \, d\tilde{M}_t + (\sigma\theta^\nu + \delta\tilde{\nu}\lambda^0 + \mu + \lambda^0\delta - r) \, dt\big)$$

where

$$\tilde{W} = W - \int \theta^\nu \, ds \qquad \text{and} \qquad \tilde{M} = M - \int \tilde{\nu} \, \lambda^0 \, ds$$

are local P^*-martingales. Hence, if we let

$$\tilde{\nu}_t = (\nu_t - \lambda^0)/\lambda^0$$

we need

$$\theta_t^\nu = -\frac{b(Y_t) - r(Y_t) + \delta(Y_t)\lambda^0(\tilde{\nu}_t + 1)}{\sigma(Y_t)} = -\frac{b(Y_t) - r(Y_t) + \delta(Y_t)\nu_t}{\sigma(Y_t)}$$

in order for $\tilde{S}$ to be a local P^*-martingale. Thus, $P^* \in \mathcal{D}$.

To see that $\mathcal{D} \subseteq \mathcal{Y}(1)$ observe that $D \in \mathcal{D}$ is a positive local martingale. An application of Itô's formula to $\tilde{S}D$ implies that this is a local martingale as well, and hence a supermartingale and this yields the claim. $\square$

In the following we will keep the notation

$$\theta^\nu = -\frac{b - r + \delta\nu}{\sigma}$$

and apply our results to two specific utility functions.

4.3 Logarithmic utility

The easiest case is as usual the logarithmic utility. As mentioned before, we will first consider the classical robust setting, i.e. a pessimistic investor and then transfer the results to an optimistic investor. We assume that the investor uses the utility function

$$U(x) = \log(x)$$

and the discount factor is equal to 1. I.e., the agent tries to maximize

$$\inf_{Q \in \mathcal{Q}} E_Q\Big[\int_0^T \log(c_t)\,\mu(dt)\Big] = \kappa \inf_{Q \in \mathcal{Q}} E_Q\Big[\int_0^T \gamma \log(c_t)\,dt + \log(c_T)\Big].$$

4.3.1 HJB equation for the dual problem

In order to find the optimal strategy we will describe the solution to the dual problem via an HJB equation. Hence, we want to apply Theorem 4.2. Therefore and since U obviously satisfies Assumption 3.2, we need to check whether Assumption 3.4 is valid.

Assumption 3.4 is satisfied if there exists $Q_0 \in \mathcal{Q}$ and $x > 0$ such that $u_{Q_0}(x) < \infty$. Now observe that $P \in \mathcal{Q}$ and that the derivation of (3.11) implies

$$u_P(x) \le E\Big[\int_0^T \tilde{U}(yD_t^\nu)\,\mu(dt)\Big] + yx$$

for all $y > 0$ and $\nu \in \mathcal{N}$. Hence, it is enough to show

$$E\Big[\int_0^T -\log(D_t^\nu)\,\mu(dt)\Big] < \infty$$

for some $\nu \in \mathcal{N}$. This will be obvious from equation (4.6) for bounded ν and thus, all assumptions necessary to apply the results of the preceding chapter are satisfied.

Let $z > 0$. Then the dual problem is according to Theorem 4.2

$$\begin{aligned}\tilde{u}(z) &= \inf_{\nu \in \mathcal{N}} \inf_{\lambda \in \Lambda} E_{Q_\lambda}\Big[\int_0^T \tilde{U}\Big(\frac{zD_t^\nu}{S_t^0 Z_t^\lambda}\Big)\,\mu(dt)\Big] \\ &= \inf_{\nu \in \mathcal{N}} \inf_{\lambda \in \Lambda} E_{Q_\lambda}\Big[\int_0^T \Big(-1 - \log\frac{zD_t^\nu}{S_t^0 Z_t^\lambda}\Big)\,\mu(dt)\Big]\end{aligned}$$

where

$$D_t^\nu = \mathcal{E}\Big(\int_0^\cdot \theta_s^\nu\,dW_s + \int_0^\cdot \frac{\nu_s - \lambda^0}{\lambda^0}\,dM_s\Big)_t, \qquad 0 \le t \le T$$

and

$$Z_t^\lambda = \mathcal{E}\Big(\int_0^{\cdot} \frac{\lambda_s - \lambda^0}{\lambda^0}\, dM_s\Big)_t, \qquad 0 \le t \le T.$$

Then we get the solution to the primal problem as

$$u(x) = \min_{z>0}(\tilde{u}(z) + zx) = \tilde{u}(1/x) + 1.$$

First, we make some computations and reduce the problem to a single minimization problem.

Lemma 4.3
We have for all $t \ge 0$

$$\inf_{\nu\in\mathcal{N}} \inf_{\lambda\in\Lambda} E_{Q_\lambda}\Big[-1 - \log \frac{D_t^\nu}{S_t^0 Z_t^\lambda}\Big] = \\ \inf_{\lambda\in\Lambda} E_{Q_\lambda}\Big[-1 + \int_0^t \frac{1}{2}(\theta_s^{\nu^*})^2 + \nu_s^* - \lambda_s + (\log\lambda_s - \log\nu_s^*)\lambda_s - r\, ds\Big]$$

where

$$\begin{aligned} \nu_s^* =& \nu^*(Y_s, \lambda_s) \\ =& \begin{cases} \frac{1}{2}\frac{b(Y_s)\delta(Y_s)+\sigma^2(Y_s)}{\delta^2(Y_s)} + \sqrt{\frac{(b(Y_s)\delta(Y_s)+\sigma(Y_s)^2)^2}{4\delta(Y_s)^4} + \frac{\sigma(Y_s)^2\lambda_s}{\delta(Y_s)^2}}, & \textit{if } \delta(Y_s) \ne 0 \\ \lambda_s, & \textit{if } \delta(Y_s) = 0. \end{cases} \end{aligned} \tag{4.4}$$

Proof. First we get

$$\begin{aligned} \frac{S_t^0 Z_t^\lambda}{D_t^\nu} &= \exp\Big\{-\int_0^t \theta_s^\nu\, dW_s + \int_0^t \Big(\frac{1}{2}(\theta_s^\nu)^2 + \nu_s - \lambda^0 + r\Big)\, ds\Big\} \exp\Big(-\int_0^t (\lambda_s - \lambda^0)\, ds\Big) \\ &\qquad \prod_{s\le t, \Delta N_s = 1} \Big(\frac{1 + \frac{\lambda_s - \lambda^0}{\lambda^0}}{1 + \frac{\nu_s - \lambda^0}{\lambda^0}}\Big) \\ &= \exp\Big\{-\int_0^t \theta_s^\nu\, dW_s + \int_0^t \Big(\frac{1}{2}(\theta_s^\nu)^2 + \nu_s - \lambda_s + r\Big)\, ds\Big\} \prod_{s\le t, \Delta N_s = 1} \frac{\lambda_s}{\nu_s}. \end{aligned} \tag{4.5}$$

Hence,

$$\begin{aligned} \log \frac{S_t^0 Z_t^\lambda}{D_t^\nu} &= -\int_0^t \theta_s^\nu\, dW_s + \sum_{s\le t}(\log\lambda_s - \log\nu_s)\Delta N_s + \int_0^t \Big(\frac{1}{2}(\theta_s^\nu)^2 + \nu_s - \lambda_s + r\Big)\, ds \\ &= -\int_0^t \theta_s^\nu\, dW_s + \int_0^t (\log\lambda_s - \log\nu_s)\, dM_s^\lambda \\ &\quad + \int_0^t \Big(\frac{1}{2}(\theta_s^\nu)^2 + \nu_s - \lambda_s + r + (\log\lambda_s - \log\nu_s)\lambda_s\Big)\, ds \end{aligned} \tag{4.6}$$

where

$$M_t^\lambda = N_t - \int_0^t \lambda_s\, ds$$

is the compensated Poisson process with respect to Q_λ. Observe that M^λ and W are Q_λ-martingales (the latter due to $[W, M] = 0$).

We would now like to take the Q_λ-expectation and argue that the integrals with respect to the martingales M^λ and W vanish such that we can minimize the remaining drift term with respect to ν. Unfortunately this approach does not work directly since the integrands are not necessarily bounded.

Instead, we will mimic the solution in [23] and reformulate our problem in terms of relative entropy. For this we will associate a measure to each $D^\nu \in \mathcal{D}$. Obviously, this is not possible on $(\Omega, \mathcal{F})$ if $E[D^\nu_T] \neq 1$.

However, if we use the notion of extended martingale measure as introduced in Föllmer, Gundel [13] the process D^ν defines a probability measure $\bar{P}_\nu$ on the probability space $(\bar{\Omega}, \bar{\mathcal{F}})$.

This probability space is defined as $\bar{\Omega} := \Omega \times (0, \infty]$ with the predictable σ-field

$$\bar{\mathcal{F}} = \sigma(\{A \times (t, \infty] \mid A \in \mathcal{F}_t, t \geq 0\})$$

(for $t > T$ define $\mathcal{F}_t = \mathcal{F}_T$). Furthermore, the filtration $(\bar{F}_t)_{t \geq 0}$ is given by

$$\bar{\mathcal{F}}_t :\ = \sigma(A \times (s, \infty] \mid A \in \mathcal{F}_s, s \leq t), \qquad t \geq 0.$$

Then a given measure Q on $(\Omega, \mathcal{F})$ is associated to the measure $\bar{Q} := Q \otimes \delta_\infty$ on $(\bar{\Omega}, \bar{\mathcal{F}})$ and for a measure $\tilde{Q}$ on $(\bar{\Omega}, \bar{\mathcal{F}})$ they define the projection Q^t on $(\Omega, \mathcal{F}_t)$ by $Q^t[A] := \bar{Q}[A \times (t, \infty]]$ for $A \in \mathcal{F}_t$. Furthermore, each $(\mathcal{F}_t)$-stopping time is associated with an $(\bar{\mathcal{F}}_t)$-stopping time via

$$\bar{\tau}(\omega, t) := \tau(\omega) \mathbb{1}_{(\tau(\omega), \infty]}(t).$$

On this measure space we can associate each $D^\nu \in \mathcal{D}$ (being a positive supermartingale) with a probability measure $\bar{P}_\nu$ (Föllmer measure) given by

$$\bar{P}_\nu[A \times (t, \infty]] = E[D^\nu_t I_A], \quad 0 \leq t < \infty, A \in \mathcal{F}_t.$$

With the above definition the set $\mathcal{Y}(1)$ can be identified with the set of extended martingale measures from [13, Definition 4.1].

Here we will use [23, Lemma 3.4] which formulates our problem as a problem for the relative entropy: Let $\nu \in \mathcal{N}$ and $Q \ll P$. If Z is defined as the density process of Q with respect to P then for any bounded $(\mathcal{F}_t)$-stopping time τ the probability measure $\bar{Q} := Q \otimes \delta_\infty$ satisfies $\bar{Q} \ll \bar{P}_\nu$ on $\bar{F}_{\bar{\tau}}$ and the relative entropy of $\bar{Q}$ with respect to $\bar{P}_\nu$ on $\bar{\mathcal{F}}_{\bar{\tau}}$ is given by

$$H_{\bar{\mathcal{F}}_{\bar{\tau}}}(\bar{Q} | \bar{P}_\nu) = E_Q\Big[\log \frac{Z_\tau}{D^\nu_\tau}\Big]. \tag{4.7}$$

We return to our calculation and apply a sequence of stopping times $(\tau_n)_{n \in \mathbb{N}}$ suitable to our problem, namely let

$$\tau_n := \inf\Big\{t \geq 0 \,\Big|\, \int_0^t \big(\nu_s^2 + |\log \nu_s|\big)\, ds \geq 1/n\Big\}.$$

Now we can stop the processes in Equation (4.6) and use that the integrals with respect to W and M^λ are martingales. Hence, we get

$$E_{Q_\lambda}\Big[-1-\log\frac{(D^\nu)^{\tau_n}_t}{(S^0)^{\tau_n}_t(Z^\lambda)^{\tau_n}_t}\Big]=$$
$$-1+E_{Q_\lambda}\Big[\int_0^{\tau_n\wedge t}\Big(\frac{1}{2}\Big((\theta^\nu)^2+\nu_s-\lambda_s+r(Y_s)+(\log\lambda_s-\log\nu_s)\lambda_s\Big)\,ds\Big]. \quad (4.8)$$

We now take the limit of this expression as $n\to\infty$ (and thus $\tau_n\nearrow\infty$). The right-hand side converges to

$$-1+E_{Q_\lambda}\Big[\int_0^t\Big(\frac{1}{2}\Big(\frac{b(Y_s)-r(Y_s)+\delta(Y_s)\nu_s}{\sigma(Y_s)}\Big)^2+\nu_s-\lambda_s+r(Y_s)$$
$$+(\log\lambda_s-\log\nu_s)\lambda_s\Big)\,ds.\Big]$$

This convergence is due to the monotone convergence theorem - the integrand is always non-negative. (To see this for $\nu_s-\lambda_s+(\log\lambda_s-\log\nu_s)\lambda_s$ observe that this expression is minimal for $\nu=\lambda$ where it equals 0.)

For the left-hand side we have

$$\sup_{n\in\mathbb{N}}E_{Q_\lambda}\Big[\log\frac{Z_{\tau_n\wedge t}}{D^\nu_{\tau_n\wedge t}}\Big]=\sup_{n\in\mathbb{N}}H_{\bar{\mathcal{F}}_{\overline{\tau_n\wedge t}}}(\bar{Q}_\lambda|\bar{P}_\nu)=H_{\bar{\mathcal{F}}_t}(\bar{Q}_\lambda|\bar{P}_\nu).$$

To get the second equality we need $\bar{\tau}_n\nearrow\infty$ and continuity properties of the relative entropy (compare e.g. [16, Proposition (15.6)]

Together with (4.7) for $\tau=t$ this yields

$$E_{Q_\lambda}\Big[-\log\frac{D^\nu_t}{S^0_tZ^\lambda_t}\Big]=$$
$$E_{Q_\lambda}\Big[\int_0^t\Big(\frac{1}{2}\Big(\frac{b(Y_s)-r(Y_s)+\delta(Y_s)\nu_s}{\sigma(Y_s)}\Big)^2+\nu_s-\lambda_s+r(Y_s)+\lambda_s\log\frac{\lambda_s}{\nu_s}\Big)\,ds.\Big] \quad (4.9)$$

Minimizing the expression in the expectation with ordinary calculus methods with respect to ν yields (under the condition that $\nu>0$) a unique minimum at

$$\nu^*_s=\begin{cases}-\frac{1}{2}\frac{(b-r)\delta+\sigma^2}{\delta^2}+\sqrt{\frac{((b-r)\delta+\sigma^2)^2}{4\delta^4}+\frac{\sigma^2\lambda_s}{\delta^2}}, & \text{if } \delta\neq 0\\ \lambda_s & \text{otherwise.}\end{cases}$$

Hence, the claim follows. □

In the following we use the notation

$$\beta=-((b-r)\delta+\sigma^2)/2\delta^2.$$

In order to use the dynamic programming principle we define

$$J(t,y,\nu,\lambda):=E\Big[\int Z^\lambda_s\log\frac{S^0_sZ^\lambda_s}{D^\nu_s}\,\tilde{\mu}_t(ds)\Big]$$

where
$$\tilde{\mu}_t(ds) = \gamma I_{[0,t]} ds + \delta_t(ds)$$
and $Y_0 = y$. Furthermore, let the value function
$$V(t, y) := \inf_{\nu \in \mathcal{N}} \inf_{\lambda \in \Lambda} J(t, y, \nu, \lambda).$$
Obviously, $\kappa \tilde{\mu}_T = \mu$ and hence,
$$\tilde{u}(z) = -1 - \log z + \kappa V(T, Y_0).$$

We can use the lemma above to guess the form of the HJB equations with the help of classical stochastic control results. For this we will need the generator (or Q-matrix) A of the Markov chain Y. Our construction of Y yields that under the measure Q_λ

$$A_t^\lambda = \begin{pmatrix} -\lambda_t & \lambda_t p_{1,2} & \dots & & \lambda_t p_{1,n} \\ & \ddots & & & \\ & & & \ddots & \\ \lambda_t p_{n,1} & \dots & & \lambda_t p_{n,n-1} & -\lambda_t \end{pmatrix}, \qquad 0 \le t \le T.$$

Observe that $(\lambda_t)_{0 \le t \le T}$ is a predictable stochastic process. Thus, it might itself depend on Y.

Next, we will prove that the following system of HJB equations characterizes the value function. The solution needs to satisfy
$$\begin{aligned} v_t^i(t) &= \inf_{\lambda \in [a_1, a_2]} \Big((1 + \gamma t) c_i(\lambda) + (A^\lambda v(t, .))_i \Big) \qquad (4.10) \\ &= \inf_{\lambda \in [a_1, a_2]} \Big((1 + \gamma t) c_i(\lambda) + \lambda \Big(\sum_{j=1}^n p_{i,j} v^j(t) - v^i(t) \Big) \Big) \end{aligned}$$
and the boundary condition
$$v^i(0) = 0 \text{ for } i = 1, \dots, n. \qquad (4.11)$$
Here, for $\delta(e_i) \neq 0$
$$\begin{aligned} c_i(\lambda) =& \frac{1}{2} \Big(\frac{b(e_i) - r(e_i) + \delta(e_i) \Big(\beta(e_i) + \sqrt{\beta(e_i)^2 + \sigma(e_i)^2 \lambda / \delta(e_i)^2} \Big)}{\sigma(e_i)} \Big)^2 - \lambda \\ & + \beta(e_i) + \sqrt{\beta(e_i)^2 + \frac{\sigma(e_i)^2 \lambda}{\delta(e_i)^2}} + r(e_i) \\ & + \Big(\log \lambda - \log \Big(\beta(e_i) + \sqrt{\beta(e_i)^2 + \frac{\sigma(e_i)^2 \lambda}{\delta(e_i)^2}} \Big) \Big) \lambda \end{aligned}$$
and for $\delta(e_i) = 0$
$$c_i(\lambda) = \frac{1}{2} \Big(\frac{b(e_i) - r(e_i)}{\sigma(e_i)} \Big)^2.$$

Theorem 4.4
There exists a unique classical solution $v \in C^1(\mathbb{R}, \mathbb{R}^n)$ to (4.10) and (4.11). This solution satisfies $v = V$. If λ^* is a measurable process, realizing the infimum in (4.10) then $\hat{\lambda} := (\lambda^*(T-t, Y_{t-}))_{0 \le t \le T}$ is a feasible control strategy from the set Λ. Furthermore, if we define $\hat{\nu}$ via (4.4) then $\hat{\nu}$ is an element of $\mathcal{N}$. Furthermore, we have then $V(t,y) = J(t,y,\hat{\nu},\hat{\lambda})$.

Proof. We will prove the result only for the case $n = 2$. The proof for the case $n > 2$ works along the same lines. Existence of a unique solution is the statement of Lemma 4.5. To prove that $v = V$ let $\nu \in \mathcal{N}$ and λ be a predictable process with values in $[a_1, a_2]$. Then Lemma 4.3 implies

$$
\begin{aligned}
J(t,y,\nu,\lambda) &= E_{Q_\lambda}\Big[\int \log \frac{S_t^0 Z_t^\lambda}{D_t^\nu}\, d\tilde{\mu}_t(s)\Big] \\
&= E_{Q_\lambda}\Big[\int \frac{1}{2}(\theta_s^\nu)^2 + \nu_s - \lambda_s + r + (\log \lambda_s - \log \nu_s)\lambda_s\, d\tilde{\mu}_t(s)\Big]
\end{aligned}
$$

which is minimal for

$$
\nu_s^* = \begin{cases} \beta + \sqrt{\beta^2 + \sigma^2\lambda_s/\delta^2} & \text{if } \delta \neq 0 \\ \lambda_s & \text{otherwise.} \end{cases}
$$

Now we show that $v \le V$. For this we choose a function $v \in C^1(\mathbb{R} \times \mathbb{R}^2, \mathbb{R})$ such that $v(t, e_i) = v^i(t)$ for $i = 1, 2$. (We name both functions v since they only differ in the way the arguments are represented.) We will also sometimes write $c(e_i, \lambda)$ instead of $c_i(\lambda)$.

Then we have according to Itô's Lemma

$$
\begin{aligned}
dv(u-t, Y_t) &= -\, v_t(u-t, Y_{t-})\, dt + v_y(u-t, Y_{t-})\, dY_t + \Delta v(u-t, Y_t) \\
&\quad -\, v_y(u-t, Y_{t-})\Delta Y_t \\
&= -\, v_t(u-t, Y_{t-})\, dt + \Delta v(u-t, Y_t)
\end{aligned}
$$

We have a closer look at the jumps of v. Since $v \in C^1$ there will only be jumps if $(N_t)_{0 \le t \le T}$ and hence $(Y_t)_{0 \le t \le T}$ jumps. If we now use $\hat{Y}_t$ to denote the state "opposite" to Y_t (i.e. $\hat{e}_1 = e_2$ and vice versa), we can write

$$
\begin{aligned}
\sum_{0 \le t \le u} \Delta v(u-t, Y_t) &= \sum_{0 \le t \le u} \big(v(u-t, \hat{Y}_{t-}) - v(u-t, Y_{t-})\big)\Delta N_t \\
&= \int_0^t \big(v(u-t, \hat{Y}_{t-}) - v(u-t, Y_{t-})\big)(dM_t^\lambda + \lambda_t\, dt).
\end{aligned}
$$

Here, M^λ is the Q_λ-martingale $N - \int \lambda_s\, ds$.

With this and the continuity of the dt-integrals in the boundary we get

$$\begin{aligned}
\int_0^u dv(u-t,Y_t) &= v(0,Y_u) - v(u,y) \\
&= \int_0^u \big(-v_t(u-t,Y_{t-}) + (v(u-t,\hat{Y}_{t-}) - v(u-t,Y_{t-}))\lambda_t\big)\,dt \\
&\qquad + \int_0^u (v(u-t,\hat{Y}_{t-}) - v(u-t,Y_{t-}))\,dM_t^\lambda \\
&= \int_0^u \big(-v_t(u-t,Y_t) + (v(u-t,\hat{Y}_t) - v(u-t,Y_t))\lambda_t\big)\,dt \\
&\qquad + \int_0^u (v(u-t,\hat{Y}_{t-}) - v(u-t,Y_{t-}))\,dM_t^\lambda \\
&\geq \int_0^u (1+\gamma(u-t))c(Y_t,\lambda_t)\,dt + \int_0^u (v(u-t,\hat{Y}_{t-}) - v(u-t,Y_{t-}))\,dM_t^\lambda \\
&= \int_0^u \Big(c(Y_t,\lambda_t) + \gamma\int_0^t c(Y_s,\lambda_s)\,ds\Big)\,dt \\
&\qquad + \int_0^u (v(u-t,\hat{Y}_{t-}) - v(u-t,Y_{t-}))\,dM_t^\lambda.
\end{aligned}$$

The inequality is a consequence of (4.10) and the last equation follows with integration by parts. Since v is bounded this implies that

$$v(u,y) \leq E_Q\left[\int_0^u \Big(c(Y_t,\lambda_t) + \gamma\int_0^t c(Y_s,\lambda_s)\,ds\Big)\,dt\right] \leq J(u,y,\lambda,\nu).$$

To show the other inequality take

$$\lambda^*(t,e_i) = \operatorname{argmin}_{\lambda\in\Lambda}\big((1+\gamma t)c_i(\lambda) + \lambda(v(t,\hat{e}_i) - v(t,e_i))\big).$$

In Lemma 4.9 we prove that c_i is a strictly convex function of λ (and it is obviously differentiable). Hence, $\lambda^*(t,e_i)$ is either a_1, a_2 or a solution to

$$c_i'(\lambda) = \frac{v(t,e_i) - v(t,\hat{e}_i)}{1+\gamma t}.$$

Since the right hand side is a continuous function of t and $c_i'(\lambda)$ is strictly increasing, $\lambda^*(t,e_i)$ is a continuous function of t.

Then $\hat{\lambda}_s = \lambda^*(u-s,Y_{s-})$ is an admissible control strategy and due to (4.10) we get $v = V$. $\square$

We will now show that the ODE (4.10) has a unique solution. In the next subsection we will consider some properties of this solution. In the following we will stick to the case $n = 2$.

Lemma 4.5
The ODE (4.10) satisfies a global Lipschitz condition. Hence, it has a unique solution.

Proof. For this we show first that for vectors $v = (v^1, v^2)$ and $w = (w^1, w^2)$ in $\mathbb{R}^2$ we have

$$\big|\,|v^1 - v^2| - |w^1 - w^2|\,\big| \le \xi \|v - w\|$$

for a fixed constant ξ. Since all norms in $\mathbb{R}^2$ are equivalent we can show this claim for the maximum norm, i.e. we take $\|v\|_{\max} = \max(v_1, v_2)$. Then we have either $(v^1 - v^2)(w^1 - w^2) \ge 0$ which yields

$$\big|\,|v^1 - v^2| - |w^1 - w^2|\,\big| = |v^1 - v^2 - (w^1 - w^2)| \le |v^1 - w^1| + |v^2 - w^2| \le 2\|v - w\|_{\max}$$

or $(v^1 - v^2)(w^1 - w^2) < 0$. In this case we can argue

$$\begin{aligned}\big|\,|v^1 - v^2| - |w^1 - w^2|\,\big| &\le \big|\,|v^1 - v^2| + |w^1 - w^2|\,\big| = |v^1 - v^2 - (w^1 - w^2)| \\ &\le |v^1 - w^1| + |v^2 - w^2| \le 2\|v - w\|_{\max}.\end{aligned}$$

We prove now the Lipschitz continuity. Let $v = (v^1, v^2)$ and $w = (w^1, w^2)$ be two vectors from $\mathbb{R}^2$. We need

$$\left\| \begin{pmatrix} \inf_\lambda((1+\gamma t)c_1(\lambda) + \lambda(v^2 - v^1)) \\ \inf_\lambda((1+\gamma t)c_2(\lambda) + \lambda(v^1 - v^2)) \end{pmatrix} - \begin{pmatrix} \inf_\lambda((1+\gamma t)c_1(\lambda) + \lambda(w^2 - w^1)) \\ \inf_\lambda((1+\gamma t)c_2(\lambda) + \lambda(w^1 - w^2)) \end{pmatrix} \right\| \le L\|v - w\|$$

for a constant L.

We assume without loss of generality that

$$\inf_\lambda (c_1(\lambda) + \lambda(v^2 - v^1)) \ge \inf_\lambda (c_1(\lambda) + \lambda(w^2 - w^1)).$$

If now λ^* is such that

$$\inf_\lambda ((1+\gamma t)c_1(\lambda) + \lambda(w^2 - w^1)) = (1+\gamma t)c_1(\lambda^*) + \lambda^*(w^2 - w^1),$$

then we have

$$\begin{aligned} 0 \le &\inf_\lambda (c_1(\lambda)(1+\gamma t) + \lambda(v^2 - v^1)) - \inf_\lambda((1+\gamma t)c_1(\lambda) + \lambda(w^2 - w^1)) \\ \le &c_1(\lambda^*)(1+\gamma t) + \lambda^*(v^2 - v^1)) - (c_1(\lambda^*)(1+\gamma t) + \lambda^*(w^2 - w^1)) \\ = &\lambda^*((v^2 - v^1) - (w^2 - w^1)) \\ \le &\lambda^* \xi \|v - w\|. \end{aligned}$$

The same argument can be applied to the second component. Since λ^* is bounded the claim follows. $\square$

Remark 4.6
If $\delta = 0$ the HJB-equation simplifies considerably to

$$v_t^i(t) = \inf_{\lambda \in [a_1, a_2]} \Big((1+\gamma t)\Big(\frac{(b(e_i) - r(e_i))^2}{2\sigma(e_i)^2} + r(e_i)\Big) + \lambda\big(v^j(t) - v^i(t)\big)\Big) \quad \text{for } (i,j) = (1,2),(2,1).$$

Especially the optimal λ will either be a_1 or a_2. Hence, in this case we have a switching control.

4.3.2 Optimistic investor

Let us now consider an optimistic investor. The problem for the optimist is

$$\max_{c \in \mathcal{A}(x)} \sup_{Q \in \mathcal{Q}} E_Q\Big[\int_0^T \log(c_t)\,\mu(dt)\Big] = \sup_{Q \in \mathcal{Q}} \max_{c \in \mathcal{A}(x)} E_Q\Big[\int_0^T \log(c_t)\,\mu(dt)\Big].$$

We will use the same notation as for the pessimist but use the index *opt* (e.g. J^{opt}).

The problem of maximizing $E_Q[\log(X_T)]$ for a given measure $Q \in \mathcal{Q}$ can be solved as before, if we assume that the investor does not face any ambiguity. Hence, we know the dual problem and the HJB equation for this problem. In order to adapt this solution for the optimist we need to maximize this solution with respect to λ. Thus, we consider the system of ODE's

$$v_t^i(t) = \sup_{\lambda}\Big((1+\gamma t)c_i(\lambda) + \lambda\big(\sum_{j=1}^n p_{i,j}v^j(t) - v^i(t)\big)\Big) \qquad \text{for } i = 1,\dots,n \tag{4.12}$$

With boundary condition $v^i(0) = 0$ for $i = 1, \dots, n$.

Then we have the following corollary.

Corollary 4.7
There exists a unique classical solution $v \in C^1(\mathbb{R}, \mathbb{R}^n)$ to (4.12). This solution satisfies $v^{opt} = V^{opt}$. If λ^* is a measurable process, realizing the supremum in (4.12) then $\lambda^{opt} := (\lambda^*(T-t, Y_{t-}))_{0 \le t \le T}$ is a feasible control strategy from the set Λ. Furthermore, ν_t^{opt} defined as in (4.4) is an element of $\mathcal{N}$, and we have then $V^{opt}(t,y) = J^{opt}(t, y, \nu^{opt}, \lambda^{opt})$.

Due to Lemma 4.9 we know that λ^* is the optimizer of a convex function. Here, we are looking for the maximum. Hence, the optimal λ^* is either a_1 or a_2 (switching control) and does not need to be unique. However, as long as $v^1 \neq v^2$ there will be at most one t where non-uniqueness might occur.

Remark 4.8
The corresponding versions of Lemma 4.5 and Remark 4.6 are also valid for the optimist. The switching controls in the case $\delta = 0$ are for pessimist and optimist polar.

4.3.3 Behavior of the solution to the HJB equation

In the following we will study the long-term behavior of the solution $v = (v^1, v^2)$ of the HJB equation (4.10) (all results will also hold for the solution to (4.12)). This behavior depends on the values of the right-hand side of (4.10) at time $t = 0$, in particular on the functions c_1, c_2. The infimum of these functions gives the slope at time 0 and this indicates how "attractive" a set of parameters is. More precisely, we will show that the value functions v^1, v^2 are always non-decreasing and that the set of parameters that is advantageous at time 0 yields a value function that dominates the other value function at all times.

We will assume $\delta \neq 0$ for this subsection.

First we give some properties of the functions c_1, c_2. We already used that these functions are convex. Hence, the minimum is unique. Furthermore, they are non-negative which yields that the value function is non-decreasing at time 0 (and we will see that it stays non-decreasing).

Lemma 4.9
The functions c_i, $i = 1, 2$ are non-negative and have a zero if and only if $-b(e_i)/\delta(e_i) > 0$ and $r(e_i) = 0$. If this zero exists it is $-b(e_i)/\delta(e_i)$. Furthermore, $c_i''(\lambda) > 0$. Hence, the term inside the infimum in (4.10) is a strictly convex function of λ.

Proof. To prove the first claim we minimize the integrand of (4.9) with respect to λ. This gives

$$c_i(\lambda) \geq 1/2(b(e_i) - r(e_i) + \delta(e_i)\nu)^2/\sigma(e_i)^2 + r(e_i) \geq 0$$

and equality for $r(e_i) = 0$ at $-b(e_i)/\delta(e_i) > 0$ (if this is non-negative).

To see the second claim observe that

$$\begin{aligned}
c_i''(\lambda) =& \frac{1}{4\lambda(\beta(e_i)^2\delta(e_i)^2 + \lambda\sigma(e_i)^2)^2}\Big(2\beta(e_i)^4\delta(e_i)^4 + 4\beta(e_i)^2\delta(e_i)^2\lambda\sigma(e_i)^2 \\
&+ 2\beta(e_i)^3\delta(e_i)^4\sqrt{\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2} + \lambda\sigma(e_i)^2\Big(2\lambda\sigma(e_i)^2 \\
&- \big(b(e_i)\delta(e_i) - r(e_i)\delta(e_i) + \sigma(e_i)^2\big)\sqrt{\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2}\Big)\Big) \\
=& \frac{\delta(e_i)^4}{2\lambda(\beta(e_i)^2\delta(e_i)^2 + \lambda\sigma(e_i)^2)^2}\Big(\big(\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2\big)^2 \\
&+ \beta(e_i)\sqrt{\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2}\big(\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2\big)\Big) \\
>& \frac{\delta(e_i)^4}{2\lambda\big(\beta(e_i)^2\delta(e_i)^2 + \lambda\sigma(e_i)^2\big)^2}\Big(\big(\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2\big)^2 \\
&- \big(\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2\big)\big(\beta(e_i)^2 + \lambda\sigma(e_i)^2/\delta(e_i)^2\big)\Big) \\
=&0.
\end{aligned}$$

$\square$

In the following proposition we consider the problem without consumption. As mentioned before we can use the functions c_1 and c_2 to decide which set of parameters yields the dominating value function. This function also has a higher slope. But as t tends to infinity the slopes converge to the same constant.

Proposition 4.10
If $\gamma = 0$ and

$$\inf_{\lambda \in [a_1, a_2]} c_1(\lambda) > \inf_{\lambda \in [a_1, a_2]} c_2(\lambda)$$

then $v_t^1(t) > v_t^2(t)$ *for all* $t \geq 0$*. Furthermore,*

$$\lim_{t\to\infty} v^1(t) - v^2(t) = \xi \quad \textit{and} \quad v_t^i \to \tilde{\xi}$$

where ξ *and* $\tilde{\xi}$ *are constants.*

Proof. At time $t = 0$ we have $v_t^1(0) > v_t^2(0)$ due to our assumption on c_1, c_2. Hence, $v^1 - v^2$ increases up to

$$t^* = \inf\{t \geq 0 \,|\, v_t^1(t^*) - v_t^2(t^*) < 0\}$$

where $t^* > 0$. Our claim is that $t^* = \infty$. Assume now, that $t^* < \infty$. Then there is some $\varepsilon > 0$ such that

$$v_t^1(t^* + \varepsilon) - v_t^2(t^* + \varepsilon) < 0.$$

Due to continuity we can choose ε and $t_0 < t^*$ such that

$$v^1(t_0) - v^2(t_0) = v^1(t^* + \varepsilon) - v^2(t^* + \varepsilon) =: \Delta v.$$

Since we have

$$v_t^1(t^* + \varepsilon) = \inf_{\lambda}(c_1(\lambda) + \lambda \Delta v) = v_t^1(t_0)$$

and similarly $v_t^2(t^* + \varepsilon) = v_t^2(t_0)$ we get

$$v_t^1(t^* + \varepsilon) - v_t^2(t^* + \varepsilon) = v_t^1(t_0) - v_t^2(t_0) > 0$$

by the definition of t^*. And this contradicts our assumption $t^* < \infty$.

The fact that $v_t^1(t) > v_t^2(t)$ for all $t \geq 0$ implies that $v^1 - v^2$ is an increasing function of t and hence, v_t^1 is decreasing and v_t^2 is increasing in t. To show that $v_t^1 - v_t^2 \to 0$ as $t \to \infty$ we assume that there is some $\varepsilon > 0$ such that $v_t^1(t) - v_t^2 \geq \varepsilon$ (due to the above this difference is decreasing). This implies $v^1(t) - v^2(t) \geq \varepsilon t \to \infty$ as $t \to \infty$. If the optimizers for the infimum at time t are $(\lambda_1^t, \lambda_2^t)$ then it follows that

$$v_t^1(t) - v_t^2(t) = c_1(\lambda_1^t) - c_2(\lambda_2^t) + (\lambda_1^t + \lambda_2^t)(v^2(t) - v^1(t)) \to -\infty$$

since c_1, c_2 and λ are bounded. This contradicts the first step of the proof and the claim follows. This implies now that $v^1(t) - v^2(t)$ converges to a constant as $t \to \infty$, which in turn implies $v_t^i(t) \to \tilde{\xi}$ for $i = 1, 2$. $\square$

In the situation with consumption the asymptotic behavior changes - we prove that the value functions are asymptotically quadratic and the difference between the slopes is linear.

Proposition 4.11
If $\gamma > 0$ *and*

$$\inf_{\lambda \in [a_1, a_2]} c_1(\lambda) > \inf_{\lambda \in [a_1, a_2]} c_2(\lambda)$$

then $v_t^1(t) > v_t^2(t)$ *for all* $t \geq 0$*. Furthermore,* $v_t^i = O(t)$*,* $\left(v^1 - v^2\right) = O(t)$ *and* $v^i = O(t^2)$ *for* $i = 1, 2$ *where* O *is the usual asymptotic notation.*

Proof. Let

$$t^* = \inf\{t \geq 0 \,|\, v_t^1(t^*) - v_t^2(t^*) < 0\}.$$

We want to prove that $t^* = \infty$. First, our assumption immediately implies $t^* > 0$. Assume that $t^* < \infty$. Then there exists $t_2 > t^*$ such that $v_t^1(t_2) - v_t^2(t_2) < 0$ and $v^1(t_2) - v^2(t_2) > 0$. The function

$$f(t) = (v^1(t) - v^2(t))/(1 + \gamma t)$$

is continuous and satisfies

$$f(0) = 0 < f(t_2) < f(t^*).$$

Hence, due to the Intermediate Value Theorem there exists $t_1 \in (0, t^*)$ such that $f(t_1) = f(t_2)$. Thus, we have

$$\begin{aligned}
v_t^1(t_2) - v_t^2(t_2) =& \inf_{\lambda \in [a_1,a_2]} \Big((1 + \gamma t_2) c_1(\lambda) + \lambda \big(v^2(t_2) - v^1(t_2) \big) \Big) \\
& - \inf_{\lambda \in [a_1,a_2]} \Big((1 + \gamma t_2) c_2(\lambda) + \lambda (v^1(t_2) - v^2(t_2)) \Big) \\
=& (1 + \gamma t_2) \Big(\inf_{\lambda \in [a_1,a_2]} \Big(c_1(\lambda) + \lambda \frac{v^2(t_2) - v^1(t_2)}{1 + \gamma t_2} \Big) \\
& - \inf_{\lambda \in [a_1,a_2]} \Big(c_2(\lambda) + \lambda \frac{v^1(t_2) - v^2(t_2)}{1 + \gamma t_2} \Big) \Big) \\
=& (1 + \gamma t_2) \Big(\inf_{\lambda \in [a_1,a_2]} \Big(c_1(\lambda) + \lambda \frac{v^2(t_1) - v^1(t_1)}{1 + \gamma t_1} \Big) \\
& - \inf_{\lambda \in [a_1,a_2]} \Big(c_2(\lambda) + \lambda \frac{v^1(t_1) - v^2(t_1)}{1 + \gamma t_1} \Big) \Big) \\
=& \frac{1 + \gamma t_2}{1 + \gamma t_1} \left(v_t^1(t_1) - v_t^2(t_1) \right) > 0.
\end{aligned}$$

This is a contradiction to the choice of t_2 and it follows $t^* = \infty$.

Now we prove $v^1(t) - v^2(t) = O(t)$. From the above we know that $v^1 - v^2 > 0$. Hence,

$$(1 + \gamma t) c_2^0 < v_t^2(t) < v_t^1(t) < (1 + \gamma t) c_1^*$$

for all $t > 0$ where

$$c_2^0 = \inf_{\lambda \in [a_1,a_2]} c_2(\lambda) \quad \text{and} \quad c_1^* = \sup_{\lambda \in [a_1,a_2]} c_1(\lambda).$$

This implies that $v_t^i = O(t)$. Hence, $v^i = O(t^2)$ and due to the form of v_t^i the difference $v^1 - v^2 = O(t)$ as well. □

Remark 4.12

A version of the above results is also valid for an optimistic investor. Here, the question which set of parameters is preferable depends on the supremum over c_i. Also the proofs differ only in the fact that sup and inf have to be exchanged.

4.3.4 Numerical Results

Using MATLAB (the code is given in the appendix) we can compute the solution of the HJB equation for given parameters σ, δ and b.

We show the results for a typical set of parameters (the parameters are given below the graphs). More precisely, we have that

$$\inf_{\lambda \in [a_1, a_2]} c_2(\lambda) > \inf_{\lambda \in [a_1, a_2]} c_1(\lambda) > 0.$$

As we proved in the preceding subsection this implies $v^2 \geq v^1$ for a pessimistic investor. Furthermore the optimizing λ for c_1 needs to be non-increasing whereas the optimizing λ for c_2 will be non-decreasing (this follows from (4.10) and $v^2 - v^1 > 0$).

First there are two graphs that give the form of $c_i(\lambda)$.

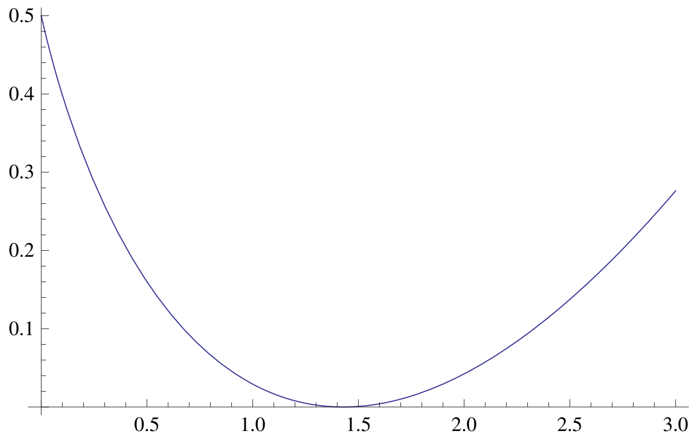

Figure 4.1: $c_1(\lambda)$ for $b = 1$, $\delta = -0.7$ and $\sigma = 1$

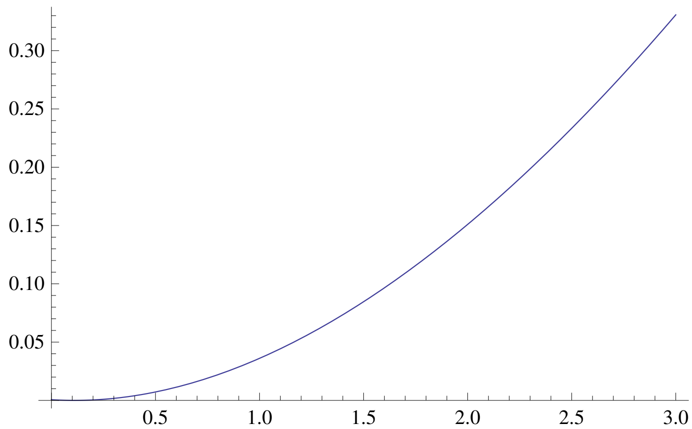

Figure 4.2: $c_2(\lambda)$ for $b = 0.035$, $\delta = -0.316$ and $\sigma = 1$

Due to Proposition 4.10 we already know that v^2, v^1 will increase linearly in the limit with the same slope. The solution for the corresponding HJB-equation without consumption ($\gamma = 0$) is shown below.

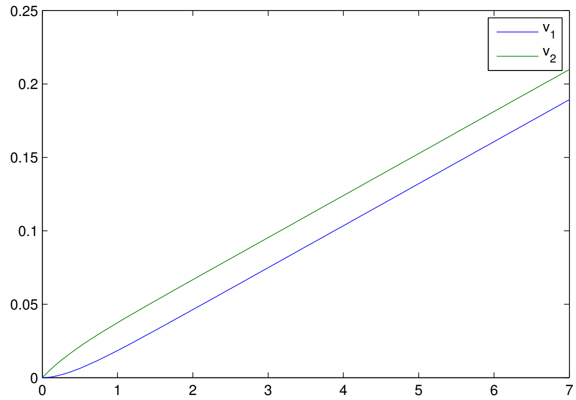

Figure 4.3: $a_1 = 1.2, a_2 = 30, T = 7$

Here the optimal λ's develop as follows.

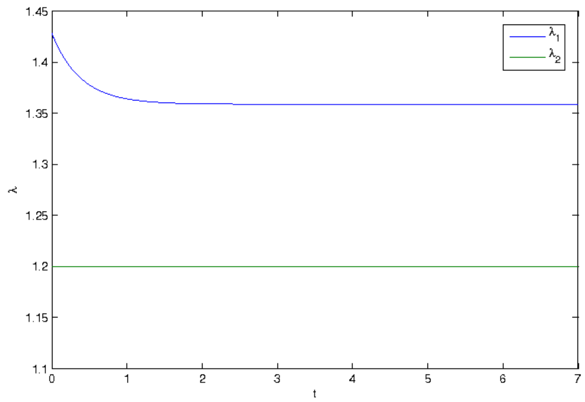

Figure 4.4: $a_1 = 1.2, a_2 = 30, T = 7$

When we include consumption ($\gamma = 1$) Proposition 4.11 implies that the solutions are asymptotically quadratic functions:

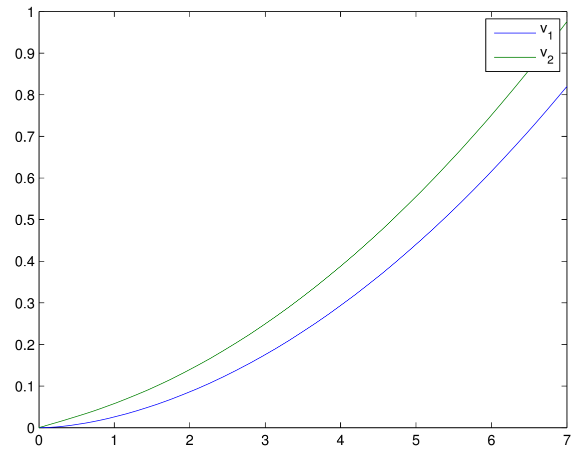

Figure 4.5: $a_1 = 1.2, a_2 = 30, T = 7$

And the optimal λ's develop as shown below.

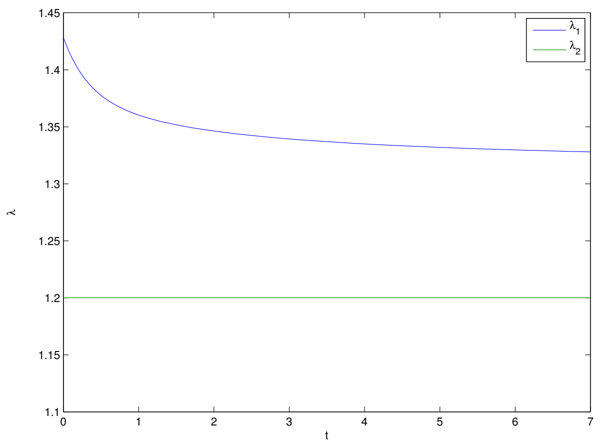

Figure 4.6: $a_1 = 1.2, a_2 = 30, T = 7$

In the following the solution for the optimist without consumption is shown. As Remark 4.12 yields it behaves basically as the solution for the pessimist. The same would be true if we included consumption. Thus, we did not include a further graph. As mentioned before, the optimal λ is always either a_1 or a_2.

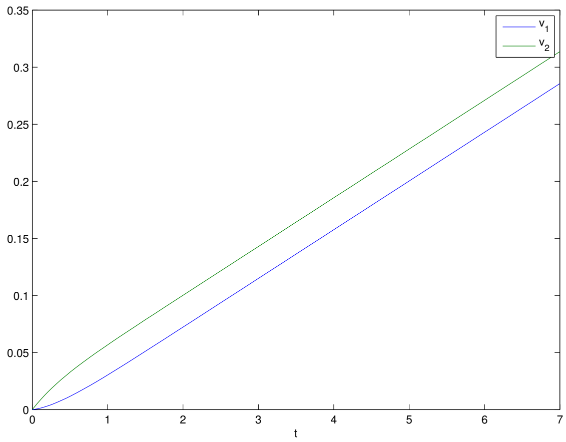

Figure 4.7: $a_1 = 1.2, a_2 = 1.5, T = 7$

And the optimal λ's develop as follows.

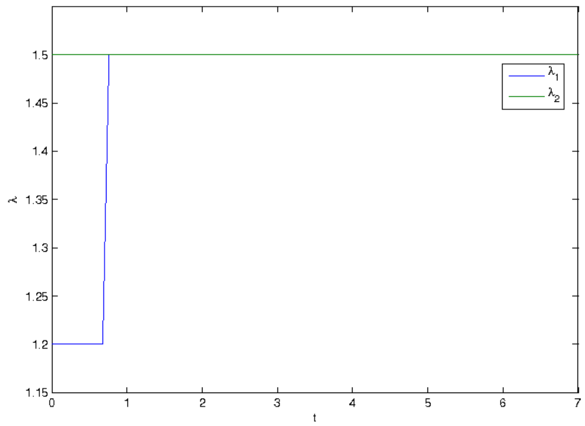

Figure 4.8: $a_1 = 1.2, a_2 = 1.5, T = 7$

4.3.5 Optimal strategy

Proposition 4.13
The optimal strategy is given as

$$\hat{\pi}(t, Y_t) = \begin{cases} -\frac{\theta_t^{\hat{\nu}}}{\sigma(Y_t)} = \frac{b(Y_t) - r(Y_t) + \hat{\nu}_t \delta(Y_t)}{\sigma(Y_t)^2} = \frac{\hat{\lambda}_t - \hat{\nu}_t}{\delta(Y_t)\hat{\nu}_t}, & \text{if } \delta(Y_t) \neq 0 \\ \frac{b(Y_t) - r(Y_t)}{\sigma(Y_t)^2}, & \text{if } \delta(Y_t) = 0, \end{cases} \qquad 0 \leq t \leq T$$

where for the pessimist $\hat{\lambda}$ and $\hat{\nu}$ are as in Theorem 4.4 and for the optimist we choose $\hat{\lambda} = \lambda^{opt}$, $\hat{\nu} = \nu^{opt}$ as in Corollary 4.7.

The optimal consumption rate is

$$\hat{c} = x\frac{Z^{\hat{\lambda}} S^0}{D^{\hat{\nu}}}$$

and the optimal wealth process is given by

$$\hat{X} = \kappa(\gamma(T - \cdot) + 1)\hat{c}.$$

The optimal consumption rate is the rate relative to μ. It implies an optimal terminal wealth of $\hat{X}_T = C_T - C_{T-} = \kappa \hat{c}_T$ and an optimal consumption rate of $\tilde{c} = \kappa\gamma\hat{c}$ with respect to the Lebesgue measure. Furthermore, the pessimist's optimal strategy is unique since the minimizing $\hat{\lambda}$ and hence $\hat{\nu}$ is unique. As mentioned before, the optimist's strategy is dt-a.e. unique.

Proof. First the form of $\hat{c}$ follows directly from Theorem 3.5 and the fact that $I(y) = 1/y$ and $\hat{z} = 1/x$.

To compute the optimal strategy we will consider the P-martingale $R = (R_t)_{0\le t\le T}$ defined as

$$R_t = \Big(\frac{\hat{X}_t}{S^0_t} + \int_0^t \frac{\hat{c}_u}{S^0_u}\,\mu(du)\Big) D^{\hat{\nu}}_t, \qquad 0 \le t \le T. \tag{4.13}$$

We have for $0 \le t \le T$

$$\begin{aligned} R_t = E[R_T|\mathcal{F}_t] &= E\Big[D^{\hat{\nu}}_T \int_0^T \frac{\hat{c}_u}{S^0_u}\,\mu(du)\,\Big|\,\mathcal{F}_t\Big] \\ &= D^{\hat{\nu}}_t \int_0^t \frac{\hat{c}_u}{S^0_u}\,\mu(du) + E\Big[D^{\hat{\nu}}_T \int_0^T \frac{x Z^{\hat{\lambda}}_u S^0_u}{D^{\hat{\nu}}_u S^0_u}\,\mu(du)\,\Big|\,\mathcal{F}_t\Big] \\ &= D^{\hat{\nu}}_t \int_0^t \frac{\hat{c}_u}{S^0_u}\,\mu(du) + x\int_t^T Z^{\hat{\lambda}}_t\,\mu(du) \\ &= D^{\hat{\nu}}_t \int_0^t \frac{\hat{c}_u}{S^0_u}\,\mu(du) + x\kappa Z^{\hat{\lambda}}_t(\gamma(T-t)+1). \end{aligned} \tag{4.14}$$

Comparison of (4.13) and (4.14) yields

$$\hat{X}_t = \kappa(\gamma(T-t)+1)\hat{c}_t \quad \text{for} \quad 0 \le t \le T.$$

Next, we can apply Itô's formula to the right-hand side and get (compare (4.5) for the form of $d\hat{c}_t$)

$$\begin{aligned} d\hat{X}_t = &- \hat{c}_t\kappa\gamma\,dt + \\ &\hat{X}_{t-}\Big(-\theta^{\hat{\nu}}_t\,dW_t + \frac{\hat{\lambda}_t - \hat{\nu}_t}{\hat{\nu}_t}\,dM_t + \big((\theta^{\hat{\nu}}_t)^2 + \hat{\nu}_t - \hat{\lambda}_t + r_t + \lambda^0\frac{\hat{\lambda}_t - \hat{\nu}_t}{\hat{\nu}_t}\big)\,dt\Big). \end{aligned}$$

Furthermore, we know that the wealth process evolves according to the SDE

$$\begin{aligned} d\hat{X}_t &= -\,\hat{c}_t\,\mu(dt) + \hat{X}_{t-}\Big(\sigma(Y_t)\pi_t\,dW_t + (b(Y_t) - r(Y_t))\pi_t + r(Y_t)\,dt + \pi_t\delta(Y_t)\,dN_t\Big) \\ &= -\,\kappa\gamma\hat{c}_t\,dt + \hat{X}_{t-}\Big(\sigma(Y_t)\pi_t\,dW_t + \pi_t\delta(Y_t)\,dM_t \\ &\qquad + \big((b(Y_t) - r(Y_t) + \lambda^0\delta(Y_t))\pi_t + r(Y_t)\big)\,dt\Big). \end{aligned} \tag{4.15}$$

Comparison of the dW_t or dM_t-terms gives the form of $\hat{\pi}$. (To see that this also implies equality for the dt-terms one has to consider the dependence of $\hat{\nu}$ from $\hat{\lambda}$.)
$\square$

4.4 Comparison of optimist and pessimist

Even though the robust problem is based on an axiomatic approach to uncertainty we interpret in this section the set of priors $\mathcal{Q}$ rather literally as set of possible market measures. (Especially for the α-MEU formulation this seems to be feasible.) In this setting we compare how the optimistic/pessimistic agent performs in a "real" market setting.

More precisely, we want to compare whether the investment strategy of the optimist or the pessimist is more successful if the "real" market measure is Q_λ for some $\lambda \in \Lambda$. First, the question arises how to evaluate the success of a strategy. One thing that seems reasonable to compare is expected terminal wealth, i.e.

$$E_{Q_\lambda}[X_T]$$

- here we will give some numerical results. However, actually one should compare expected utility

$$E_{Q_\lambda}[\log(X_T)]$$

since the agents choose the strategy such that the expected utility is maximized.

Except from the case $\delta = 0$ we are only able to give numerical results. Then it is obviously not possible to simulate *all* possible market measures, hence we restrict ourselves to measures that are given by a constant $\lambda \in \Lambda$. Furthermore, we will consider only terminal wealth, no consumption. It would be interesting to do more advanced simulations in particular to have a non constant intensity λ.

First, we observe that for the case $\delta = 0$ the expected utility from terminal wealth does not depend on the measure the investor used to optimize. I.e. optimist and pessimist have the same expected utility, namely

$$E_Q[\log X_T] = E_Q\Big[\int_0^T (1/2(b-r)^2/\sigma^2 + r)\, ds\Big]$$

(the equation follows with easy computations from Proposition 4.13). In this case both investors have as optimal λ's a_1 or a_2 but they get polar values (compare Remark 4.8).

For the more general case we used MATLAB to simulate the outcome. (The code for the simulation is again given in the appendix.) Since we assume that the "real" market has a constant intensity we choose the parameters for the simulation such that optimist and pessimist optimize also with respect to (nearly) constant intensities.

For the parameters $T = 0.55$, $\delta = (-1, -1)$, $\sigma = (0.1, 1)$, $b = (0.5, 0.54)$, $a_1 = 0.8$, $a_2 = 1.1$ we have that the λ for the pessimist is constant at 0.8 and λ for the optimist is constant at 1.1. We get the following results for the expected utility

λ	0.8	0.9	1	1.1
optimist	0.0162	0.0588	0.1010	0.1399
pessimist	0.0412	0.0655	0.0899	0.1162

and the expected wealth

λ	0.8	0.9	1	1.1
optimist	1.2122	1.2870	1.3754	1.4617
pessimist	1.1013	1.1390	1.1722	1.2108

Next, we consider the parameters $T = 1.1$, $\delta = (-0.25, -0.25)$, $\sigma = (0.7, 0.177)$, $b = (1, 0.5)$ and $x = 1$. Then the λ for the pessimist is higher (not constant, but always above 1.8) and $\lambda^{opt} = 0.1$ (at least up to T). Here the results for the expected utility are

λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7
optimist	1.0444	0.9793	0.8961	0.8288	0.7496	0.6754	0.6024
pessimist	0.6982	0.6727	0.6479	0.6175	0.5962	0.5708	0.5423
λ	0.8	0.9	1	1.1	1.2	1.3	1.4
optimist	0.5324	0.4499	0.3830	0.3087	0.2358	0.1675	0.0964
pessimist	0.5197	0.4894	0.4728	0.4414	0.4181	0.3912	0.3666
λ	1.5	1.6	1.7	1.8	1.9		
optimist	0.0190	-0.0510	-0.1238	-0.1961	-0.2652		
pessimist	0.3390	0.3190	0.2902	0.2676	0.2429		

and

λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7
optimist	8.1494	7.5340	7.1730	6.8569	6.5238	6.1784	5.7942
pessimist	2.4095	2.3603	2.3130	2.2529	2.2047	2.1587	2.1006
λ	0.8	0.9	1	1.1	1.2	1.3	1.4
optimist	5.6687	5.2179	5.0162	4.6934	4.5034	4.2550	4.0486
pessimist	2.0588	2.0178	1.9693	1.9186	1.8779	1.8352	1.8019
λ	1.5	1.6	1.7	1.8	1.9		
optimist	3.7914	3.6007	3.3875	3.2082	3.1150		
pessimist	1.7591	1.7175	1.6814	1.6395	1.6082		

for the expected wealth.

As is clear from the problem formulation the optimist performs better than the pessimist in his "predicted" scenario and vice versa. Furthermore, the "optimistic" scenario gives the higher utility for both agents and the utility is monotonic in λ. Apart from that the main difference between the two approaches is that the range of possible expected utilities is smaller for the pessimistic agent. This is in accordance with the "pessimistic" view. This investor prefers safety over possible gains whereas the optimistic investor bids on the uncertainty and might thus gain and lose more. With respect to expected wealth the optimist always gets more in our examples.

4.5 HARA utility

We will now consider our problem for an investor with HARA utility function with risk aversion parameter $0 < \alpha < 1$, i.e. $U(x) = x^\alpha/\alpha$. We will see later that a negative risk aversion parameter is not feasible for the HJB equations we are computing. We will again develop the corresponding HJB equation which enables the investor to compute the optimal strategy (at least numerically).

4.5.1 HJB equation

The aim of the pessimistic investor is again to maximize

$$\inf_{Q\in\mathcal{Q}} E_Q\Big[\int_0^T U(c_t)\,\mu(dt)\Big]$$

where now

$$\mu(dt) = \kappa(\gamma e^{-\rho t}dt + \delta_{\{T\}}(dt)).$$

We give again the HJB equation for the dual problem. As before we want to use Theorem 4.2. For this observe that U obviously satisfies Assumption 3.2 and Assumption 3.4 is fulfilled with the same argument as for the logarithmic utility when we use equation (4.20).

Hence, we have that the dual problem is

$$\tilde{u}(z) = \inf_{\lambda\in\Lambda}\inf_{\nu\in\mathcal{N}} E\Big[\,Z_T^\lambda\int_0^T \tilde{U}\Big(\frac{zD_t^\nu}{S_t^0 Z_t^\lambda}\Big)\,\mu(dt)\,\Big], \qquad z > 0.$$

Here $\tilde{U}(z) = -z^\beta/\beta$ with $\beta = -\alpha/(1-\alpha)$. In particular β will have a different meaning from the last section.

Then we know that

$$\begin{aligned} u(x) &= \min_{z>0}(\tilde{u}(z) + zx) \\ &= \min_{z>0}\Big(\inf_{\lambda\in\Lambda}\inf_{\nu\in\mathcal{N}} E\Big[\,Z_T^\lambda\int_0^T \tilde{U}\Big(\frac{zD_t^\nu}{S_t^0 Z_t^\lambda}\Big)\,\mu(dt)\,\Big] + zx\Big) \\ &= \inf_{\lambda\in\Lambda}\inf_{\nu\in\mathcal{N}}\min_{z>0}(-\frac{z^\beta}{\beta}\Gamma^{\lambda,\nu} + zx) \\ &= \frac{x^\alpha}{\alpha}(\inf_{\lambda\in\Lambda}\inf_{\nu\in\mathcal{N}}\Gamma^{\lambda,\nu})^{1-\alpha} \end{aligned} \tag{4.16}$$

where

$$\Gamma^{\lambda,\nu} = E\Big[\int_0^T Z_t^\lambda (D_t^\nu/(S_t^0 Z_t^\lambda))^\beta\,\mu(dt)\Big].$$

The last equality follows by standard optimization methods which also yield that the optimal

$$\hat{z} = (x/\Gamma^{\lambda,\nu})^{1/(\beta-1)}. \tag{4.17}$$

To simplify this expression we compute

$$
\begin{aligned}
&(Z_t^{\lambda})^{1-\beta}(D_t^{\nu})^{\beta}/(S_t^0)^{\beta} = \\
&\qquad \exp\Big(-\int_0^t (1-\beta)(\lambda_s-\lambda^0)\,ds\Big)\prod_{s\le t}\Big(1+\frac{\lambda_s-\lambda^0}{\lambda^0}\Delta N_s\Big)^{1-\beta} \\
&\qquad \exp\Big(\int_0^t \beta\,\theta_s^{\nu}\,dW_s - \frac{1}{2}\int_0^t \beta\,(\theta_s^{\nu})^2\,ds - \int_0^t \beta\,(\nu_s-\lambda^0)\,ds\Big) \\
&\qquad \prod_{s\le t}\Big(1+\frac{\nu_s-\lambda^0}{\lambda^0}\Delta N_s\Big)^{\beta}\exp\Big(-\int_0^t \beta\, r_s\,ds\Big) \\
&\qquad = \exp\Big(\int_0^t \lambda^0-\lambda_s+\beta\,(\lambda_s-\lambda^0)-\frac{1}{2}\beta\,(\theta_s^{\nu})^2-\beta\,(\nu_s-\lambda^0+r_s)\,ds\Big) \\
&\qquad \exp\Big(\int_0^t \beta\,\theta_s^{\nu}\,dW_s\Big)\prod_{s\le t,\Delta N_s\neq 0}\Big(\frac{\lambda_s}{\lambda^0}\Big)^{1-\beta}\Big(\frac{\nu_s}{\lambda^0}\Big)^{\beta} \\
&\qquad = \exp\Big(\int_0^t \beta\,\theta_s^{\nu}\,dW_s - \frac{1}{2}\int_0^t (\beta\,\theta_s^{\nu})^2\,ds\Big) \\
&\qquad \exp\Big(\int_0^t \big(\lambda^0-\lambda_s+\beta\,\big(\lambda_s-\frac{1}{2}(\theta_s^{\nu})^2-\nu_s-r_s\big)+\frac{1}{2}(\beta\,\theta_s^{\nu})^2\big)\,ds\Big) \\
&\qquad \prod_{s\le t,\Delta N_s\neq 0}\Big(\frac{\lambda_s^{1-\beta}\nu_s^{\beta}-\lambda^0}{\lambda^0}+1\Big) \\
&\qquad =\mathcal{E}\Big(\int_0^{\cdot} \beta\,\theta_s^{\nu}\,dW_s\Big)_t\,\mathcal{E}\Big(\int_0^{\cdot} \frac{\lambda_s^{1-\beta}\nu_s^{\beta}-\lambda^0}{\lambda^0}\,dM_s\Big)_t\exp\Big(\int_0^t \lambda^0-\lambda_s\,ds\Big) \\
&\qquad \exp\Big(\int_0^t \beta\,\big(\lambda_s-\frac{1}{2}(\theta_s^{\nu})^2-\nu_s-r_s\big)+\frac{1}{2}\beta^2(\theta_s^{\nu})^2+\frac{\lambda_s^{1-\beta}\nu_s^{\beta}-\lambda^0}{\lambda^0}\lambda^0\,ds\Big).
\end{aligned} \tag{4.18}
$$

We want to exploit the fact that the stochastic exponential defines a density as long as it satisfies

$$
E\Big[\mathcal{E}\Big(\int_0^{\cdot} \beta\,\theta_s^{\nu}\,dW_s+\int_0^{\cdot} \frac{\lambda_s^{1-\beta}\nu_s^{\beta}-\lambda^0}{\lambda^0}\,dM_s\Big)_T\Big]=1. \tag{4.19}
$$

In this case let

$$
\frac{dP^{\lambda,\nu}}{dP}=\mathcal{E}\Big(\int_0^{\cdot} \beta\,\theta_s^{\nu}\,dW_s+\int_0^{\cdot} \frac{\lambda_s^{1-\beta}\nu_s^{\beta}-\lambda^0}{\lambda^0}\,dM_s\Big)_T.
$$

Then

$$
E\big[(Z_t^{\lambda})^{1-\beta}(D_t^{\nu})^{\beta}/(S_t^0)^{\beta}\big]=E^{\lambda,\nu}\Big[\exp\Big(\int_0^t q(Y_s,\lambda_s,\nu_s)\,ds\Big)\Big] \tag{4.20}
$$

where

$$
q(e_i,\lambda,\nu)=-\lambda+\beta\,\big(\lambda-\frac{1}{2}(\theta^{\nu}(e_i))^2-\nu-r(e_i)\big)+\frac{1}{2}\beta^2(\theta^{\nu}(e_i))^2+\lambda^{1-\beta}\nu^{\beta}.
$$

Observe that (4.19) is satisfied if

$$\nu \in \mathcal{N}_0 := \{\nu \in \mathcal{N} \mid \int_0^T \nu_s^2 \, ds \text{ is bounded } P\text{-a.s.}\}.$$

This follows by an application of Novikov's criterion and [35, remark, p. 142]. In analogy to [22, Lemma 3.2] we can formulate the following Lemma.

Lemma 4.14
For fixed $\lambda \in \Lambda$ we have

$$\inf_{\nu \in \mathcal{N}} \Gamma^{\nu,\lambda} = \inf_{\nu \in \mathcal{N}_0} \Gamma^{\nu,\lambda}. \tag{4.21}$$

Proof. The proof is a copy of the proof of [22, Lemma 3.2] if we use Lemma 3.12 instead of [40, Lemma 3.6]. □

In the end, we need to compute (4.16). The above lemma implies that it is enough to compute

$$\inf_{\lambda \in \Lambda, \nu \in \mathcal{N}_0} \Gamma^{\lambda,\nu} = \inf_{\lambda \in \Lambda, \nu \in \mathcal{N}_0} E^{\lambda,\nu} \Big[\int_0^T \exp\Big(\int_0^u q(Y_s, \lambda_s, \nu_s) \, ds \Big) \, \mu(du) \Big].$$

We will again use stochastic programming to find ordinary differential equations that give the solution. For this define for $0 \le t \le T$, $\nu \in \mathcal{N}_0$, $\lambda \in \Lambda$ and $i \in \{1, \ldots, n\}$

$$J(t, e_i, \nu, \lambda) = E^{\lambda,\nu} \Big[\int \exp\Big(\int_0^u q(Y_s, \lambda_s, \nu_s) \, ds \Big) \, \tilde{\mu}_t(du) \Big]$$

where

$$\tilde{\mu}_t(du) = \tilde{\kappa} e^{\rho(t-u)} I_{[0,t]}(du) + \delta_t(du),$$

$\tilde{\kappa} = \gamma e^{-\rho T}$ and $Y_0 = e_i$. Hence,

$$J(T, y, \nu, \lambda) = 1/\kappa \, \Gamma^{\lambda,\nu}.$$

Furthermore, let

$$V(t, y) = \inf_{\nu \in \mathcal{N}_0, \lambda \in \Lambda} J(t, y, \nu, \lambda).$$

If the generator of the Markov chain with respect to the measure $P^{\lambda,\nu}$ is called $A^{\lambda,\nu}$ the HJB's should be

$$\begin{aligned} v_t^i(t) &= \inf_{\lambda \in [a_1,a_2], \nu > 0} \big(q(e_i, \lambda, \nu) v^i(t) + (A^{\lambda,\nu} v(t,.))_i \big) + \tilde{\kappa} e^{-\rho t} \\ &= \inf_{\lambda \in [a_1,a_2], \nu > 0} \big(q(e_i, \lambda, \nu) v^i(t) + \lambda^{1-\beta} \nu^{\beta} \sum_{j=1}^n p_{i,j} (v^j(t) - v^i(t)) \big) + \tilde{\kappa} e^{-\rho t} \end{aligned} \tag{4.22}$$

with boundary condition

$$v^i(0) = 1, \qquad i = 1, \ldots, n. \tag{4.23}$$

Remark 4.15
If $\alpha < 0$ there exists no feasible solution to the HJB equation as can be seen by the following argumentation. To compute $v_t^i(0)$ we need to minimize $q(e_i, \lambda, \nu)$ with respect to ν and λ. Since λ is bounded we can here neglect the terms which only depend on λ. Hence, we need to minimize

$$\frac{1}{2}\beta(1-\beta)\Big(\frac{b-r+\delta\nu}{\sigma}\Big)^2 - \beta\nu + \lambda^{1-\beta}\nu^\beta$$

with respect to $\nu > 0$. If we consider the case $\alpha < 0$ we have $0 < \beta < 1$ and in this case q goes to minus infinity as ν goes to infinity. Hence, there exists no feasible solution to the HJB equation. Furthermore, it is not easy to prove existence for the solutions to the above HJB equations and numerical results suggest that there might be cases where no bounded solution exists.

Theorem 4.16 (Verification)
Assume that a bounded solution $v \in C^1(\mathbb{R}, \mathbb{R}^n)$ to (4.22), (4.23) exists, then $v = V$. If (λ^*, ν^*) is a measurable bounded process, realizing the infimum in (4.22) then $(\hat{\lambda}, \hat{\nu}) := (\lambda^*(T-t, Y_{t-}), \nu^*(T-t, Y_{t-}))_{0 \le t \le T}$ is a feasible control strategy from the set $\Lambda \times \mathcal{N}_0$. Furthermore, we have then $V(t, y) = J(t, y, \hat{\nu}, \hat{\lambda})$.

Proof. As for the logarithmic utility we will prove the result only for $n = 2$ in order to keep the notation easy.

First, we show that $v \le V$. As in the proof for the logarithmic utility let $v : \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function which satisfies $v(t, e_i) = v^i(t)$ for $i = 1, 2$. Note, that since v is differentiable $v(t, Y_t)$ can only jump when Y jumps. For fixed $\lambda \in \Lambda, \nu \in \mathcal{N}_0$ Itô's lemma yields

$$\begin{aligned}
&d(e^{\int_0^t q(Y_s,\lambda_s,\nu_s)\,ds}\, v(u-t, Y_t)) \\
&\quad = e^{\int_0^t q(Y_s,\lambda_s,\nu_s)\,ds}\Big(\big(q(Y_{t-}, \lambda_t, \nu_t)\, v(u-t, Y_{t-}) - v_t(u-t, Y_{t-})\big)\,dt + \Delta v(u-t, Y_t)\Big) \\
&\quad \ge e^{\int_0^t q(Y_s,\lambda_s,\nu_s)\,ds}\Big((v(u-t, \hat{Y}_{t-}) - v(u-t, Y_{t-}))\,dM_t^{\lambda,\nu} - \tilde{\kappa} e^{-\rho(u-t)}\,dt\Big)
\end{aligned}$$

where

$$M^{\lambda,\nu} = N - \int_0^{\cdot} \lambda_s^{1-\beta}\nu_s^\beta\,ds$$

is the $P^{\lambda,\nu}$-compensated Poisson process and the inequality follows with (4.22). With the same arguments as for the logarithmic utility the result does not change if we use t instead of $t-$. And thus, we get by integrating and taking expectations

$$\begin{aligned}
v(u, y) &\le E^{\lambda,\nu}\Big[e^{\int_0^u q(Y_s,\lambda_s,\nu_s)\,ds} + \int_0^u e^{\int_0^t q(Y_s,\lambda_s,\nu_s)\,ds}\tilde{\kappa} e^{-\rho(u-t)}\,dt\Big] \\
&= J(u, y, \nu, \lambda).
\end{aligned}$$

To see the other inequality we take again for $i \ne j$

$$\begin{aligned}
(\lambda^*(t, e_i), \nu^*(t, e_i)) = \operatorname{argmin}_{(\lambda,\nu)\in[a_1,a_2]\times(0,\infty)}\big(&q(e_i, \lambda, \nu)v^i(t) \\
&+ \lambda^{1-\beta}\nu^\beta(v^j(t) - v^i(t)) + \tilde{\kappa}e^{-\rho t}\big)
\end{aligned}$$

Arguing as for the logarithmic utility and with the help of Lemma 4.17 we get that (λ^*, ν^*) can be chosen as a measurable function.

Then $(\hat{\lambda}_s, \hat{\nu}_s) = (\lambda^*(T-s, Y_{s-}), \nu^*(T-s, Y_{s-}))$ is an admissible control strategy and due to (4.22) we get $v = V$. □

In the following we stick to the case where the Markov chain switches only between two states.

4.5.2 Behavior of the solution to the HJB equation

As for the logarithmic utility we can again try to find properties of the solution to the HJB equation. Here, the situation is not as easy as before since we have the additional infimum over ν. However, we can still prove that the "better" state of the Markov chain can be determined at time 0. For this we consider first the function q and show that it is non-negative and convex in ν.

Lemma 4.17
We have $q(e_i, \lambda, \nu) \geq 0$ and $q_{\nu\nu}(e_i, \lambda, \nu) > 0$ for all $i \in \{1, 2\}$, $\lambda \in [a_1, a_2]$ and $\nu > 0$.

Proof. If we minimize q for given $\nu > 0$ we get that the minimizing λ equals ν. (Obviously this λ is not necessarily in $[a_1, a_2]$.) And thus, we have

$$q(e_i, \lambda, \nu) \geq q(e_i, \nu, \nu) \geq -\nu + \beta\nu - \beta\nu + \nu^{1-\beta}\nu^{\beta} = 0.$$

Furthermore, by taking derivatives we get

$$\begin{aligned} q_\nu =& \beta\theta^\nu \frac{\delta}{\sigma} - \beta - \beta^2\theta^\nu \frac{\delta}{\sigma} + \beta\lambda^{1-\beta}\nu^{\beta-1} \qquad \text{and} \\ q_{\nu\nu} =& (\beta-1)\beta\Big(\frac{\delta^2}{\sigma^2} + \lambda^{1-\beta}\nu^{\beta-2}\Big) > 0. \end{aligned}$$

□

As before we can determine whether the state e_1 or e_2 is more advantageous by considering the infimum of the corresponding function q. This infimum gives the slope of the value function at time zero and the following proposition states that the state with the higher slope at time zero is always "preferable".

Proposition 4.18
If

$$\inf_{\nu>0} \inf_{\lambda\in[a_1,a_2]} q(e_1, \lambda, \nu) > \inf_{\nu>0} \inf_{\lambda\in[a_1,a_2]} q(e_2, \lambda, \nu)$$

then it follows

$$v_t^1(t) \geq v_t^2(t) \text{ for all } t \geq 0.$$

Proof. Let as in the corresponding proof for the logarithmic utility

$$t^* = \inf\{t \geq 0 \,|\, v_t^1(t^*) - v_t^2(t^*) < 0\}.$$

We want to prove that $t^* = \infty$. As before, our assumption immediately implies $t^* > 0$. Up to t^* we have

$$v^1_t(t) > v^2_t(t) > 0. \tag{4.24}$$

Assume that $t^* < \infty$. Then there exists also due to (4.24) $t_2 > t^*$ such that $v^1_t(t_2) - v^2_t(t_2) < 0$, $v^1(t_2) - v^2(t_2) > 0$, $v^1_t(t_2) > 0$ and $v^1(t_2) > 0$. Using this we get that the function

$$f : [0, \infty) \to \mathbb{R}, f(t) = (v^1(t) - v^2(t))/v^1(t) = 1 - v^2(t)/v^1(t)$$

is continuous and satisfies

$$f(0) = 0 < f(t_2) < f(t^*).$$

Hence, due to the Intermediate Value Theorem there exists $t_1 \in (0, t^*)$ such that $f(t_1) = f(t_2)$ and hence,

$$v^2(t_2)/v^1(t_2) = v^2(t_1)/v^1(t_1)$$

or equivalently

$$v^2(t_2)/v^2(t_1) = v^1(t_2)/v^1(t_1).$$

Together this yields

$$\begin{aligned}
v^1_t(t_2) - v^2_t(t_2) = &\inf_{\lambda\in[a_1,a_2],\nu>0}\Big(q(e_1,\lambda,\nu)v^1(t_2) + \lambda^{1-\beta}\nu^\beta\big(v^2(t_2) - v^1(t_2)\big)\Big) \\
&- \inf_{\lambda\in[a_1,a_2],\nu>0}\Big(q(e_2,\lambda,\nu)v^2(t_2) + \lambda^{1-\beta}\nu^\beta(v^1(t_2) - v^2(t_2))\Big) \\
=&v^1(t_2)\inf_{\lambda\in[a_1,a_2],\nu>0}\Big(q(e_1,\lambda,\nu) + \lambda^{1-\beta}\nu^\beta\frac{v^2(t_2) - v^1(t_2)}{v^1(t_2)}\Big) \\
&- v^2(t_2)\inf_{\lambda\in[a_1,a_2],\nu>0}\Big(q(e_2,\lambda,\nu) + \lambda^{1-\beta}\nu^\beta\frac{v^1(t_2) - v^2(t_2)}{v^2(t_2)}\Big) \\
=&\frac{v^1(t_2)}{v^1(t_1)}v^1_t(t_1) - \frac{v^2(t_2)}{v^2(t_1)}v^2_t(t_1) \\
=&\frac{v^1(t_2)}{v^1(t_1)}\big(v^1_t(t_1) - v^2_t(t_1)\big) > 0.
\end{aligned}$$

And this contradicts our assumption $t^* < \infty$. □

4.5.3 Numerical Results

We use again MATLAB to compute the solution of the HJB equation for given parameters σ, δ, r and b. We consider the utility function $U(x) = 1/2\sqrt{x}$, i.e. $\beta = -1/2$. First there are two graphs (from mathematica) that give the form of $q(\lambda, \nu)$. We choose $a_1 = 0.1$ and $a_2 = 1.5$ which according to Proposition 4.18 implies that $v^2 \geq v^1$.

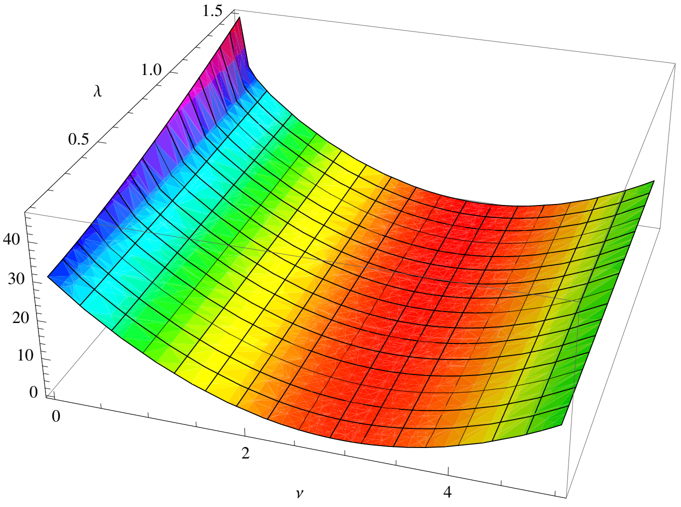

Figure 4.9: $q(\lambda, \nu)$ for $b = 0.9$, $\delta = -0.3$, $r = 0$ and $\sigma = 0.1$

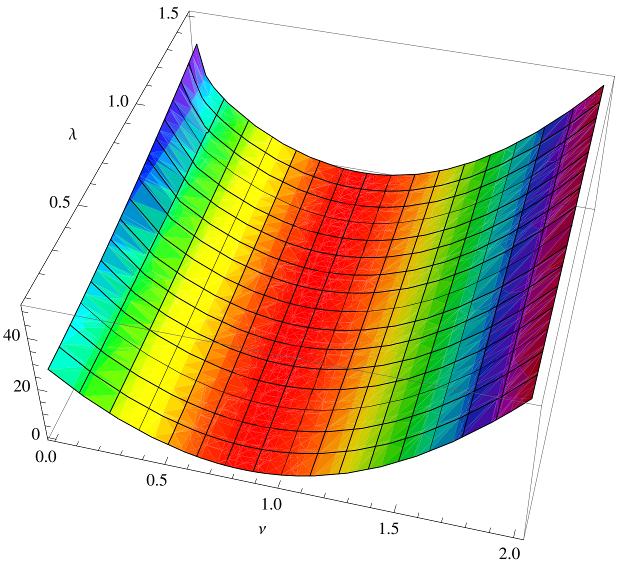

Figure 4.10: $q(\lambda, \nu)$ for $b = -0.84$, $\delta = 1$, $r = 0$ and $\sigma = 0.1$

Then the solution for the corresponding HJB-equation without consumption is given by the following graph.

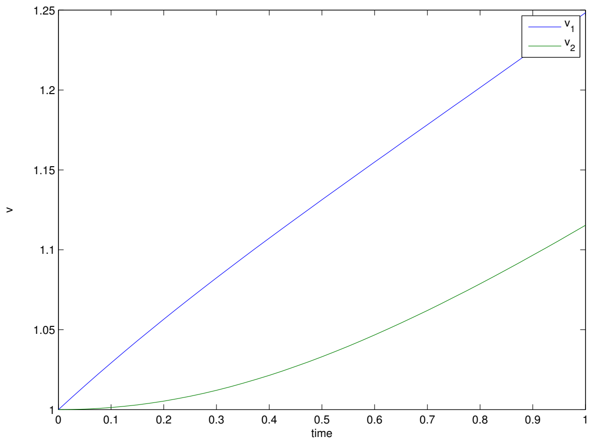

Figure 4.11: $a_1 = 0.1, a_2 = 1.5, T = 1$

Here the λ's develop as follows.

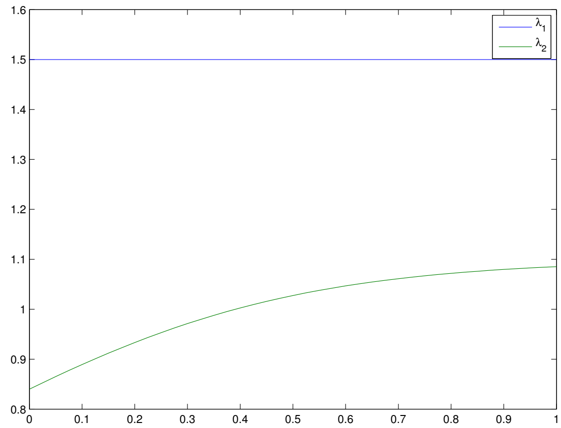

Figure 4.12: $a_1 = 0.1, a_2 = 1.5, T = 1$

4.5.4 Optimal strategy

In this subsection we will compute the optimal consumption rate and the optimal strategy.

Proposition 4.19
Assume that a solution to the HJB equation exists. Then the optimal consumption rate with respect to μ is

$$\hat{c} = \hat{z}^{\beta-1}(D^{\hat{\nu}})^{\beta-1}(Z^{\hat{\lambda}})^{1-\beta}(S^0)^{1-\beta}$$

where $\hat\lambda$ and $\hat\nu$ are as in Theorem 4.16 and

$$\hat z = \Big(\frac{x}{\kappa v(T, Y_0)}\Big)^{1/(\beta-1)}.$$

We get this rate by investing according to the strategy

$$\hat\pi(t, Y_t) = \frac{-\theta^{\hat\nu}_t}{\sigma(Y_t)}(1-\beta) = 1/\delta(Y_t)\Big(-1 + \frac{v(T-t, \hat Y_{t-})}{v(T-t, Y_{t-})}\Big(\frac{\hat\lambda_t}{\hat\nu_t}\Big)^{1-\beta}\Big), \quad 0 \le t \le T.$$

Proof. First, the form of $\hat c$ follows from Theorem 3.5, and we already computed

$$\hat z = \Big(\frac{x}{\kappa v(T, Y_0)}\Big)^{1/(\beta-1)}.$$

To compute the optimal strategy we will again consider the P-martingale $(R_t)_{0\le t\le T}$ defined as

$$R_t = \Big(\frac{\hat X_t}{S^0_t} + \int_0^t \frac{\hat c_u}{S^0_u}\,\mu(du)\Big)D^{\hat\nu}_t, \qquad 0 \le t \le T. \tag{4.25}$$

We use the martingale property to compute

$$R_t = E[R_T|\mathcal{F}_t] = D^{\hat\nu}_t \int_0^t \frac{\hat c_u}{S^0_u}\,\mu(du) + \hat z^{\beta-1}\Big(\frac{D^{\hat\nu}_t}{S^0_t}\Big)^{\beta}(Z^{\hat\lambda}_t)^{1-\beta}\mathcal{E}_t$$

where

$$\mathcal{E}_t = E\Big[\int_t^T \Big(\frac{D^{\hat\nu}_u}{D^{\hat\nu}_t}\Big)^{\beta}\Big(\frac{Z^{\hat\lambda}_u}{Z^{\hat\lambda}_t}\Big)^{1-\beta}\Big(\frac{S^0_u}{S^0_t}\Big)^{-\beta}\mu(du)\,\Big|\,\mathcal{F}_t\Big].$$

We introduce the controls $\hat\lambda^{(t)}_s = \lambda^*(T-t-s, Y_{s-})$ and $\hat\nu^{(t)}_s = \nu^*(T-t-s, Y_{s-})$ and use the Markov-property (compare [39]) to compute furthermore

$$\mathcal{E}_t = \kappa J(T-t, Y_t, \hat\lambda^{(t)}, \hat\nu^{(t)}) = \kappa v(T-t, Y_t).$$

Hence,

$$R_t = D^{\hat\nu}_t \int_0^t \frac{\hat c_u}{S^0_u}\,\mu(du) + \frac{x}{v(T, Y_0)}\Big(\frac{D^{\hat\nu}_t}{S^0_t}\Big)^{\beta}(Z^{\hat\lambda}_t)^{1-\beta}v(T-t, Y_t). \tag{4.26}$$

Comparison of (4.25) and (4.26) yields

$$\hat X_t = \kappa\, v(T-t, Y_t)\,\hat c_t.$$

We want to apply Itô's formula to the right-hand side of the above equation. For this we have first (the computations are the same as for (4.18))

$$\begin{aligned} d\hat c_t =& \hat c_{t-}\Big(\theta^{\hat\nu}_t(\beta-1)\,dW_t + (\hat\nu_t^{\beta-1}\hat\lambda_t^{1-\beta} - 1)\,dM_t + \Big(\frac{1}{2}(\theta^{\hat\nu}(\beta-1))^2 - \lambda^0 - \hat\lambda_t + \\ & \lambda^0\hat\nu_t^{\beta-1}\hat\lambda_t^{1-\beta} + \beta\hat\lambda_t - \beta\hat\nu_t - \frac{1}{2}(\beta-1)(\theta^{\hat\nu}_t)^2 + \hat\nu_t - \beta r + r\Big)\,dt\Big). \end{aligned}$$

Thus, we get

$$
\begin{aligned}
d\hat{X}_t =& \hat{X}_{t-}\Big(\theta_t^{\hat{\nu}}(\beta-1)\,dW_t + (\hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta}-1)\,dM_t + \Big(\frac{1}{2}(\theta_t^{\hat{\nu}}(\beta-1))^2 - \lambda^0 + \\
&\lambda^0\hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta} - \hat{\lambda}_t + \beta\hat{\lambda}_t - \beta\hat{\nu}_t - \frac{1}{2}(\beta-1)(\theta_t^{\hat{\nu}})^2 + \hat{\nu}_t - \beta r + r\Big)\,dt\Big) \\
&- \kappa v_t(T-t,Y_t)\,\hat{c}_t\,dt + \kappa(v(T-t,\hat{Y}_t) - v(T-t,Y_t))\,\hat{c}_t\,dN_t \\
&+ \kappa\,\hat{c}_t(v(T-t,\hat{Y}_t) - v(T-t,Y_t))(\hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta}-1)\,dN_t \\
=& \hat{X}_{t-}\Big(\theta_t^{\hat{\nu}}(\beta-1)\,dW_t + \big(\hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta}\frac{v(T-t,\hat{Y}_t)}{v(T-t,Y_t)} - 1\big)\,dM_t + \Big(\lambda^0\hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta} \\
&+ \frac{1}{2}(\theta_t^{\hat{\nu}}(\beta-1))^2 - \lambda^0 - \hat{\lambda}_t + \beta\hat{\lambda}_t - \beta\hat{\nu}_t - \frac{1}{2}(\beta-1)(\theta_t^{\hat{\nu}})^2 + \hat{\nu}_t - \beta r \\
&+ r\Big)\,dt\Big) - \kappa\Big(\big(-\hat{\lambda}_t + \beta\,\big(\hat{\lambda}_t - \frac{1}{2}(\theta_t^{\hat{\nu}})^2 - \hat{\nu}_t - r\big) + \frac{1}{2}\beta^2(\theta_t^{\hat{\nu}})^2 + \hat{\lambda}_t^{1-\beta}\hat{\nu}_t^{\beta}\big) \\
&v(T-t,Y_t) + \hat{\lambda}_t^{1-\beta}\hat{\nu}_t^{\beta}(v(T-t,\hat{Y}_t) - v(T-t,Y_t)) + \gamma e^{-\rho T}e^{\rho(T-t)}\Big)\,\hat{c}_t\,dt \\
&+ \kappa(v(T-t,\hat{Y}_t) - v(T-t,Y_t))\,\lambda^0\hat{c}_t(1 + \hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta} - 1)\,dt \\
=& \hat{X}_{t-}\Big(\theta_t^{\hat{\nu}}(\beta-1)\,dW_t + \big(\hat{\nu}_t^{\beta-1}\hat{\lambda}_t^{1-\beta}\frac{v(T-t,\hat{Y}_t)}{v(T-t,Y_t)} - 1\big)\,dM_t + \Big((\theta_t^{\hat{\nu}})^2(1-\beta) \\
&- \lambda^0 + \hat{\nu}_t + r + \frac{V(T-t,\hat{Y}_t)}{V(T-t,Y_t)}(\lambda^0 - \hat{\lambda}_t^{1-\beta}\hat{\nu}_t^{\beta} + \lambda^0\hat{\lambda}_t^{1-\beta}\hat{\nu}_t^{\beta-1})\Big)\,dt\Big) \\
&- \kappa\gamma e^{-\rho t}\hat{c}_t\,dt.
\end{aligned} \tag{4.27}
$$

We know again that the wealth process evolves according to the SDE

$$
\begin{aligned}
d\hat{X}_t =& -\hat{c}_t\,\mu(dt) + \hat{X}_{t-}\Big(\sigma(Y_t)\pi(t,Y_t)\,dW_t + \big((b(Y_t) - r(Y_t))\pi(t,Y_t) \\
&+ r(Y_t)\big)\,dt + \pi(t,Y_t)\delta(Y_t)\,dN_t\Big) \\
=& -\kappa\gamma e^{-\rho t}\hat{c}_t\,dt + \hat{X}_{t-}\Big(\sigma(Y_t)\pi(t,Y_t)\,dW_t + \big((b(Y_t) - r(Y_t) + \lambda^0\delta)\pi(t,Y_t) \\
&+ r(Y_t)\big)\,dt + \pi(t,Y_t)\delta(Y_t)\,dM_t\Big).
\end{aligned} \tag{4.28}
$$

We can now compare the dW and dM terms in (4.27) and (4.28) and get

$$
\hat{\pi}(t,Y_t) = \frac{-\theta_t^{\hat{\nu}}}{\sigma(Y_t)}(1-\beta) = 1/\delta(Y_t)\Big(-1 + \frac{v(T-t,\hat{Y}_{t-})}{v(T-t,Y_{t-})}\Big(\frac{\hat{\lambda}_t}{\hat{\nu}_t}\Big)^{1-\beta}\Big).
$$

(Using this strategy the dt terms are the same.) $\square$

Appendix
Matlab Code

In the following we give some code that was used for the simulations.

This program is used to compare the expected terminal wealth and logarithmic utility of an optimistic and pessimistic investor. As arguments it gets the interval for the intensity of the Poisson process, i.e. a_1, a_2. All other parameters are specified directly in the program.

```
%calls the function expectedterminalutility
% and expectedterminalwealth for values of lambda between a1
%and a2


function optpesscomparision(a1,a2)

    T=0.55; %terminal time
    delta=[-1,-1];
    sigm=[0.1,1];
    b=[0.5,0.54];
    x=1;
    %let the hjb solver run in order to get the lambda for optimist and
    %pessimist
    lambdaopt=hjblambdalogconsopt2(T,a1,a2,delta,b,sigm);
    lambdapess=hjblambdalogconsumption(T,a1,a2,delta,b,sigm);

    for l=a1:0.1:a2
        expwealth=expectedterminalwealth(l,lambdaopt,lambdapess,T,delta,sigm,b,x)
        exputility=expectedterminalutility(l,lambdaopt,lambdapess,T,delta,sigm,b,x)
    end
end
```

The above program calls *hjblambdalogconsumption.m* which solves the HJB equation for a pessimistic investor and the given parameters. In particular it returns the times at which the solution to the ODE was evaluated and the corresponding optimal values for λ. (We do not give *hjblambdalogconsopt2.m* here since it is basically identical code just for the optimistic investor.)

```
function lambdapess = hjblambdalogconsumption(T,a11,a21,delta,b,sigm)
%returns vector with vector for lambda1 and lambda2 and the timepoints for
```

```
%these

%solves the hjb equation for the logarithmic utility

clear y a1 a2;

global a1 a2 zaehl BETA1 BETA2 b1 b2 sigma1 sigma2 Y delta1
    delta2 tpoints gamma

gamma=0;
a1=a11;
a2=a21;
zaehl=1;
lambdatil1=zeros(1,10);
lambdatil2=zeros(1,10);
delta1=delta(1);
delta2=delta(2);
b1=b(1);
b2=b(2);
sigma1=sigm(1);
sigma2=sigm(2);

BETA1=-(b1*delta1+sigma1^2)/(2*delta1^2);
BETA2=-(b2*delta2+sigma2^2)/(2*delta2^2);


tspan = [0 T];
y0 = [0; 0];
% Solve the problem using ode45
[tpoints,Y]=ode45(@f,tspan,y0);
%plots the solution
plot(tpoints,Y);
lambdapess=zeros(3,length(Y));
for zaehl=1:length(Y)
      lambdatil1(zaehl)=fminbnd(@tomin21,a1,a2);
      lambdatil2(zaehl)=fminbnd(@tomin22,a1,a2);
      lambdapess(1,zaehl)=lambdatil1(zaehl);
      lambdapess(2,zaehl)=lambdatil2(zaehl);
      lambdapess(3,zaehl)=tpoints(zaehl);
end;

figure;
plot(tpoints,lambdatil1,tpoints,lambdatil2);
title('lambda for pessimist')
end
```

```
% ----------------------------------------------------------------
function dydt = f(t,w)
global a1 a2 BETA1 BETA2 b1 b2 sigma1 sigma2 Y delta1 delta2 gamma
y=w;

lambda1=fminbnd(@tomin1,a1,a2);
lambda2=fminbnd(@tomin2,a1,a2);

dydt=zeros(2,1);
dydt(1)=(1+gamma*t)*(0.5*((b1+delta1*(BETA1+sqrt(BETA1^2+sigma1^2*lambda1/delta1^2)))/sigma1)^2+BETA1+sqrt(BETA1^2+sigma1^2*lambda1/delta1^2)-lambda1+(log(lambda1)-log(BETA1+sqrt(BETA1^2+sigma1^2*lambda1/delta1^2)))*lambda1)+lambda1*(y(2)-y(1));
dydt(2)=(1+gamma*t)*(0.5*((b2+delta2*(BETA2+sqrt(BETA2^2+sigma2^2*lambda2/delta2^2)))/sigma2)^2+BETA2+sqrt(BETA2^2+sigma2^2*lambda2/delta2^2)-lambda2+(log(lambda2)-log(BETA2+sqrt(BETA2^2+sigma2^2*lambda2/delta2^2)))*lambda2)+lambda2*(y(1)-y(2));



function tomin = tomin1(l)


tomin=(1+gamma*t)*0.5*((b1+delta1*(BETA1+sqrt(BETA1^2+sigma1^2*l/delta1^2)))/sigma1)^2+BETA1+sqrt(BETA1^2+sigma1^2*l/delta1^2)-l+(log(l)-log(BETA1+sqrt(BETA1^2+sigma1^2*l/delta1^2)))*l+l*(y(2)-y(1));
end

function tomin = tomin2(l)
tomin=(1+gamma*t)*0.5*((b2+delta2*(BETA2+sqrt(BETA2^2+sigma2^2*l/delta2^2)))/sigma2)^2+BETA2+sqrt(BETA2^2+sigma2^2*l/delta2^2)-l+(log(l)-log(BETA2+sqrt(BETA2^2+sigma2^2*l/delta2^2)))*l+l*(y(1)-y(2));
end
end



function tomin = tomin22(l)
global BETA2 b2 sigma2 zaehl delta2 Y tpoints gamma
tomin=(1+gamma*tpoints(zaehl))*0.5*((b2+delta2*(BETA2+sqrt(BETA2^2+sigma2^2*l/delta2^2)))/sigma2)^2+BETA2+sqrt(BETA2^2+sigma2^2*l/delta2^2)-l+(log(l)-log(BETA2+sqrt(BETA2^2+sigma2^2*l/delta2^2)))*l+l*(Y(zaehl,1)-Y(zaehl,2));
```

```
end

function tomin = tomin21(l)
global BETA1 b1 sigma1 zaehl delta1 Y tpoints gamma
tomin=(1+gamma*tpoints(zaehl))*0.5*((b1+delta1*(BETA1+sqrt(BETA1^2+sigma1^2*l/delta1^2)))/sigma1)^2+BETA1+sqrt(BETA1^2+sigma1^2*l/delta1^2)-l+(log(l)-log(BETA1+sqrt(BETA1^2+sigma1^2*l/delta1^2)))*l+l*(Y(zaehl,2)-Y(zaehl,1));
end
```

The programs *expectedterminalutility.m* and *expectedterminalwealth.m* are also called by *optpesscomparision.m*. They simulate the terminal utility/wealth of an investor that is using a strategy adapted to a pessimistic/optimistic attitude but invests in a market where the stock price process moves according to a “real world” λ as given by the calling method.

```
%the function simulates the stock price process for a given Poisson
%intensity lambda
%in order to compute the expected terminal utility for an investor with
%optimistic/pessimistic attitude.
%the SDE for the stock price process is simulated with the Euler method. The
%timepoints are the ones used by matlab for the HJB solution. This is
%certainly one of the major reasons for numerical lack of precision since
%the distance between these points increases with time

%input: lambda: the lambda that should be used to simulate the stock price
%        lambdaopt: array with the argmin of the solution to the ODE for the
%        optimist
%        lambdapess: the same for the pessimist
%        T: terminal time
%        delta, sigm, b, x: parameters for the SDE to generate the stock
%        price
%output: an array S with expected utilities where S(1) is the expected
%utility of the pessimist and S(2) is the expected utility of the optmist


function Ex =expectedterminalutility(lambda,lambdaopt,lambdapess,T,delta,sigm,b,x)

```

```
    %number of iterations used for the Monte-Carlo-Method
    numberofiter=100000;

    S(1)=0;
    S(2)=0;


    %use the lambda to compute the nu (for Y_t=1, Y_t=2)
    nupess(1,:)=-1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2 +sqrt((1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2)^2+sigm(1)^2*lambdapess(1,:)/delta(1)^2);
    nupess(2,:)=-1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2 +sqrt((1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2)^2+sigm(2)^2*lambdapess(2,:)/delta(2)^2);
    nuopt(1,:)=-1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2 +sqrt((1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2)^2+sigm(1)^2*lambdaopt(1,:)/delta(1)^2);
    nuopt(2,:)=-1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2 +sqrt((1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2)^2+sigm(2)^2*lambdaopt(2,:)/delta(2)^2);

    %compute the strategies for different cases of Y
    pipess(1,:)=(lambdapess(1,:)-nupess(1,:))./(delta(1)* nupess(1,:));
    pipess(2,:)=(lambdapess(2,:)-nupess(2,:))./(delta(2)* nupess(2,:));
    piopt(1,:)=(lambdaopt(1,:)-nuopt(1,:))./(delta(1)* nuopt(1,:));
    piopt(2,:)=(lambdaopt(2,:)-nuopt(2,:))./(delta(2)* nuopt(2,:));

    %compute expectation for pessimist
    for(i=1:numberofiter)
        y=1;
        int1=0;
        int2=0;
        prod=1;

        timepoints=lambdapess(3,:);
        N=length(timepoints);
        dW=BM(timepoints);
        dN=PP(timepoints,lambda);
        for(tind=2:N)
            if(dN(tind)==1)
                %product changes
                prod=prod*(pipess(y,tind)*delta(y)+1);
```

```
                %switch y
                y=mod(y+1,2)+1;

            end
            int1=int1+sigm(y)*pipess(y,tind)*dW(tind);
            int2=int2+(-0.5*sigm(y)^2*pipess(y,tind)^2+b(y)*
                pipess(y,tind))*(timepoints(tind)-timepoints(
                tind-1));
        end
        S(1)=S(1)+(int1+int2)+log(prod);
    end


    %compute expectation for optimist
    for(i=1:numberofiter)
        y=1;
        int1=0;
        int2=0;
        prod=1;

        timepoints=lambdaopt(3,:);
        N=length(timepoints);
        dW=BM(timepoints);
        dN=PP(timepoints,lambda);
        for(tind=2:N)
            if(dN(tind)==1)
                prod=prod*(piopt(y,tind)*delta(y)+1);
                %switch y
                y=mod(y+1,2)+1;
                %product changes
            end
            int1=int1+sigm(y)*piopt(y,tind)*dW(tind);
            int2=int2+(-0.5*sigm(y)^2*piopt(y,tind)^2+b(y)*piopt
                (y,tind))*(timepoints(tind)-timepoints(tind-1));
        end
       S(2)=S(2)+(int1+int2)+log(prod);
    end
    Ex=S./numberofiter;
end
```

```
%the function simulates the stock price process for a given
    Poisson intensity lambda
%in order to compute the expected terminal wealth for an
    investor with
%optimistic/pessimistic attitude.
%the SDE for the stock price process is solved with the Euler
    method. The
```

```
%timepoints are the ones used by matlab for the HJB solution. This is
%certainly one of the major reasons for numerical lack of precision since
%the distance between these points increases with the time

%input: the lambda that should be used to simulate the stock price
%output: an array S with expected wealth where S(1) is the expected
%wealth of the pessimist and S(2) is the expected wealth of the optmist




function Ex =expectedterminalwealth(lambda,lambdaopt,lambdapess,T,delta,sigm,b,x)
  numberofiter=100000;

  S(1)=0;
  S(2)=0;


  %use the lambda to compute the nu (for Y_t=1, Y_t=2)
  nupess(1,:)=-1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2 +sqrt((1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2)^2+sigm(1)^2*lambdapess(1,:)/delta(1)^2);
  nupess(2,:)=-1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2 +sqrt((1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2)^2+sigm(2)^2*lambdapess(2,:)/delta(2)^2);
  nuopt(1,:)=-1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2 +sqrt((1/2*(b(1)*delta(1)+sigm(1)^2)/delta(1)^2)^2+sigm(1)^2*lambdaopt(1,:)/delta(1)^2);
  nuopt(2,:)=-1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2 +sqrt((1/2*(b(2)*delta(2)+sigm(2)^2)/delta(2)^2)^2+sigm(2)^2*lambdaopt(2,:)/delta(2)^2);
  %compute the strategies for different cases of Y
  pipess(1,:)=(lambdapess(1,:)-nupess(1,:))./(delta(1)* nupess(1,:));
  pipess(2,:)=(lambdapess(2,:)-nupess(2,:))./(delta(2)* nupess(2,:));
  piopt(1,:)=(lambdaopt(1,:)-nuopt(1,:))./(delta(1)* nuopt(1,:));
  piopt(2,:)=(lambdaopt(2,:)-nuopt(2,:))./(delta(2)* nuopt(2,:));

```

```
%compute expectation for pessimist
for(i=1:numberofiter)
    y=1;
    int1=0;
    int2=0;
    prod=1;

    timepoints=lambdapess(3,:);
    N=length(timepoints);
    dW=BM(timepoints);
    dN=PP(timepoints,lambda);
    for(tind=2:N)
        if(dN(tind)==1)
            %product changes
            prod=prod*(pipess(y,tind)*delta(y)+1);
            %switch y
            y=mod(y+1,2)+1;

        end
        %theta=-(b(y)+delta(y)*nuopt(y,tind))/sigm(y);
        int1=int1+sigm(y)*pipess(y,tind)*dW(tind);
        int2=int2+(-0.5*sigm(y)^2*pipess(y,tind)^2+b(y)*
            pipess(y,tind))*(timepoints(tind)-timepoints(
            tind-1));

    end
    S(1)=S(1)+x*exp(int1+int2)*prod;
end

%compute expectation for optimist
for(i=1:numberofiter)
    y=1;
    int1=0;
    int2=0;
    prod=1;

    timepoints=lambdaopt(3,:);
    N=length(timepoints);
    dW=BM(timepoints);
    dN=PP(timepoints,lambda);
    for(tind=2:N)
        if(dN(tind)==1)
            prod=prod*(piopt(y,tind)*delta(y)+1);
            %switch y
            y=mod(y+1,2)+1;
            %product changes
        end
```

```
                %theta=-(b(y)+delta(y)*nuopt(y,tind))/sigm(y);
                int1=int1+sigm(y)*piopt(y,tind)*dW(tind);
                int2=int2+(-0.5*sigm(y)^2*piopt(y,tind)^2+b(y)*piopt
                    (y,tind))*(timepoints(tind)-timepoints(tind-1));
            end
            S(2)=S(2)+x*exp(int1+int2)*prod;
        end
        Ex=S./numberofiter;
end
```

The following two files are used to simulate the increments of a Brownian motion/ a Poisson process as they are needed in the above simulation.

```
%generates the increments of a path
%of Brownian motion, where the timeintervals
%are given as the array timepoints

function dW=BM(timepoints)

dW=zeros(1,length(timepoints));
for i=2:length(timepoints)
    dW(i)=randn*sqrt(timepoints(i)-timepoints(i-1));
end;
end
```

```
%constructs a Poisson process for the parameter l, which is
    given at the
%for each t out of timepoints
%function returns a vector of 0 and 1, one if a jump occurs at
    the
%corresponding timepoint


function dN=PP(timepoints,l)
    N=length(timepoints);
    PP(1)=0;
    dN(1)=0;
    i=2;
    t=0;
    while (t<=timepoints(N))
        t=t+log(1-rand)/(-l);
        while (i<=N && (timepoints(i) < t))
           PP(i)=PP(i-1);
           dN(i)=0;
           i=i+1;
        end
        if(i<=N)
           PP(i)=PP(i-1)+1;
           dN(i)=1;
```

```
            i=i+1;
        end
    end
end
```

To compute the solution to the HARA utility maximization problem we use the program *hjbsolverhara.m*.

```
function hjbsolverhara(T,a1,b1)
%solves the hjb equation for the HARA utility
%a1 and b1 give the interval for lambda
clear Y a b;
global a b delta1 delta2 MU1 MU2 r1 r2 sigma1 sigma2 BETA
    tpoints Y zaehl
a=a1; %lower boundary for lambda
b=b1;    %upper boudary for lambda
BETA=-1/5;  %=-alpha/1-alpha

%parameters for stock price process
delta1=-1;
delta2=-1;
MU1=2;
MU2=2.5;
r1=1;
r2=1.1;
sigma1=.1;
sigma2=.1;

%compute the solution Y for the ode dydt and boundary condition
    Y(0)=1
tspan = [0 T];
y0 = [1; 1];
[tpoints,Y]=ode45(@f,tspan,y0);
plot(tpoints,Y);

%to plot the development of lambda and nu, compute these with
    the given
%solution Y
lambda1=zeros(1,10);
lambda2=zeros(1,10);
para(1)=(a+b)/2;
para(2)=1;
for zaehl=1:length(Y)
     para1=fminsearch(@tomin12,para);
     para2=fminsearch(@tomin22,para);
    lambda1(zaehl)=para1(1);
     lambda2(zaehl)=para2(1);
end;

```

```
figure;
plot(tpoints,lambda1,tpoints,lambda2);

end

% ------------------------------------------------------------------------
function dydt = f(t,w)
global a b delta1 delta2 MU1 MU2 r1 r2 sigma1 sigma2 BETA



%fminsearch, define function as -infty otherwise...
para(1)=(a+b)/2;
para(2)=1;
para1=fminsearch(@tomin1,para);
para2=fminsearch(@tomin2,para);
lambda1=para1(1)
nu1=para1(2);
lambda2=para2(1);
nu2=para2(2);


dydt = [ qOfY1(lambda1,nu1)*w(1)+(lambda1^(1-BETA)*nu1^BETA)*(w
    (2)-w(1))
         qOfY2(lambda2,nu2)*w(2)+(lambda2^(1-BETA)*nu2^BETA)*(w(1)
             -w(2))]

function tomin = tomin1(ln)
if (a<=ln(1) && ln(1)<=b && ln(2)>0)
    tomin=qOfY1(ln(1),ln(2))*w(1)-ln(1)^(1-BETA)*ln(2)^BETA*(w
        (2)-w(1));
else tomin=inf;
end
end

function  tomin = tomin2(ln)
if (a<=ln(1) && ln(1)<=b && ln(2)>0)
    tomin=qOfY2(ln(1),ln(2))*w(2)-ln(1)^(1-BETA)*ln(2)^BETA*(w
        (1)-w(2));
else tomin=inf;
end
end

    function qOfY1 = qOfY1(l,n)
        theta1=-(MU1-r1+delta1 *n)/sigma1;
        qOfY1=-l+BETA*(l-1/2 *theta1^2-n-r1)+1/2 *BETA^2 *theta1
            ^2+l^(1-BETA)*n^BETA;
```

```
        end
        function qOfY2 = qOfY2(l,n)
            theta2=-(MU2-r2+delta2 *n)/sigma2;
            qOfY2=-l+BETA*(l-1/2 *theta2^2-n-r2)+1/2 *BETA^2 *theta2
                ^2+ l^(1-BETA)*n^BETA;
        end
end



function tomin = tomin12(ln)
global a b Y BETA MU1 delta1 sigma1 r1 zaehl

if (a<=ln(1) && ln(1)<=b && ln(2)>0)
      tomin=qOfY1(ln(1),ln(2))*Y(zaehl,1)-ln(1)^(1-BETA)*ln(2)^
          BETA*(Y(zaehl,2)-Y(zaehl,1));
else tomin=inf;
end
    function qOfY1 = qOfY1(l,n)
        theta1=-(MU1-r1+delta1 *n)/sigma1;
        qOfY1=-l+BETA*(l-1/2 *theta1^2-n-r1)+1/2*BETA^2 *theta1
            ^2+l^(1-BETA)*n^BETA;
    end

end


function  tomin = tomin22(ln)
global a b Y BETA MU2 delta2 sigma2 zaehl r2
if (a<=ln(1) && ln(1)<=b && ln(2)>0)
    tomin=qOfY2(ln(1),ln(2))*Y(zaehl,2)-ln(1)^(1-BETA)*ln(2)^
        BETA*(Y(zaehl,1)-Y(zaehl,2));
else tomin=inf;
end

function qOfY2 = qOfY2(l,n)
            theta2=-(MU2-r2+delta2 *n)/sigma2;
            qOfY2=-l+BETA*(l-1/2 *theta2^2-n-r2)+1/2*BETA^2 *theta2
                ^2+ l^(1-BETA)*n^BETA;
        end

end
```

Bibliography

[1] J.-P. Aubin and I. Ekeland. *Applied nonlinear analysis*, chapter 6. John Wiley & Sons, New York, 1984.

[2] N. Bäuerle and U. Rieder. Portfolio optimization with jumps and unobservable intensity process. *Mathematical Finance*, 2(17):205–224, 2007.

[3] D. Becherer. Bounded solutions to backward SDE's with jumps for utility optimization and indifference hedging. *Annals of Applied Probability*, 2006.

[4] G. Bordigoni, A. Matoussi, and M. Schweizer. A stochastic control approach to a robust utility maximization problem. In *Proceedings of Abel Symposium 2005*. Springer.

[5] C. Burgert and L. Rüschendorf. Optimal consumption strategies under model uncertainty. *Statistics & Decisions*, 2004.

[6] J. Cvitanić, W. Schachermayer, and H. Wang. Utility maximization in incomplete markets with random endowment. *Finance & Stochastics*, 5(2):259–272, 2001.

[7] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(3):463–520, 1994.

[8] F. Delbaen and W. Schachermayer. *The mathematics of arbitrage.* Springer, 2006.

[9] D. Duffie and T. Zariphopoulou. Optimal investment with undiversifiable income risk. *Mathematical Finance*, 3:135–148, 1993.

[10] R. J. Elliott. *Stochastic Calculus and Applications.* Springer-Verlag, Berlin, Heidelberg, New York, 1982.

[11] D. Ellsberg. Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

[12] W. Fei. Optimal portfolio choice based on α-meu under ambiguity. *Stochastic Models*, 25(3):455–482, 2009.

[13] H. Föllmer and A. Gundel. Robust projections in the class of martingale measures. *Illinois Journal of Mathematics*, 2006.

[14] H. Föllmer and D. Kramkov. Optional decompositions under constraints. *Probability Theory and Related Fields*, 109(1):1–25, 1997.

[15] H. Föllmer and A. Schied. *Stochastic Finance. An Introduction in Discrete Time.* de Gryuter, 2nd edition, 2004.

[16] H. O. Georgii. *Gibbs measures and phase transitions.* Linkde Gruyter, 1988.

[17] P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133–173, 2004.

[18] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, (18):141–153, 1989.

[19] A. Gundel. Robust utility maximization for complete and incomplete market models. *Finance and Stochastics*, 9(2):151–176, 2005.

[20] L. P. Hansen and T. J. Sargent. Robust control and model uncertainty. *American Economic Review*, 91:60–66, 2001.

[21] D. Hernández-Hernández and A. Schied. A control approach to robust utility maximization with logarithmic utility and time consistent penalties. submitted, 2006.

[22] D. Hernández-Hernández and A. Schied. Robust utility maximization in a stochastic factor model. *Statistics & Decisions*, 24(1):107–129, 2006.

[23] D. Hernández-Hernández and A. Schied. A control approach to robust utility maximization with logarithmic utility and time consistent penalties. *Stochastic Processes and Their Applications*, 117(8):980–1000, 2007.

[24] D. Hernández-Hernández and A. Schied. Robust maximization of consumption with logarithmic utility. In *Proceedings of the 2007 American Control Conference*, pages 1120–1123, 2007.

[25] J. Hugonnier and D. Kramkov. Optimal investment with random endowments in incomplete markets. *Annals of Applied Probability*, 14(2):845–864, 2004.

[26] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, 2nd edition, 2002.

[27] I. Karatzas and G. Žitković. Optimal consumption from investment and random endowment in incomplete semimartingale markets. *Annals of Probability*, 31(4):1821–1858, 2003.

[28] D. Kramkov and W. Schachermayer. The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, 9(3):904–950, 1999.

[29] D. Kramkov and W. Schachermayer. Necessary and sufficient conditions in the problem of optimal investment in incomplete markets. *Annals of Applied Probability*, 13(4):1504–1516, 2003.

[30] V. Krätschmer. Robust representation of convex risk measures by probability measures. *Finance and Stochastics*, 9(4):597–608, 2005.

[31] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447 – 1498, 2006.

[32] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 1969.

[33] R. C. Merton. Optimum consumption and portfolio rules in a continuous time model. *Journal of Economic Theory*, 3:373–413, 1971.

[34] M. Müller. *Market completion and robust utility maximization*. PhD thesis, Humboldt University, 2005.

[35] P. E. Protter. *Stochastic integration and differential equations*. Springer, 2. ed. edition, 2004.

[36] M.-C. Quenez. Optimal portfolio in a multiple-priors model. In *Seminar on stochastic analysis, random fields and applications IV*, Progress in Probability 58, pages 291–321, Basel, 2004. Birkhäuser.

[37] A. Schied. On the Neyman-Pearson problem for law-invariant risk measures and robust utility functionals. *Annals of Applied Probability*, 14:1398–1423, 2004.

[38] A. Schied. Optimal investments for risk- and ambiguity-averse preferences: a duality approach. *Finance and Stochastics*, 11(1):107–129, 2007.

[39] A. Schied. Robust optimal control for a consumption-investment problem. *Mathematical Methods of Operations Research*, 67(1):1–20, 2008.

[40] A. Schied and C.-T. Wu. Duality theory for optimal investments under model uncertainty. *Statistics & Decisions*, 23(3):199–217, 2005.

[41] W. Wittmüss. Robust optimization of consumption with random endowment. *Stochastics An International Journal of Probability and Stochastic Processes*, 80(5):459–475, 2008.